Journal of
Experimental Psychology

VOL. 32, NO. 4 APRIL, 1943

OBSERVATIONS ON SEQUENCES OF CHOICES MADE AT FIVE SUCCESSIVE CHOICE POINTS

BY MALCOLM G. PRESTON AND PEARL M. ZEID

University of Pennsylvania

Two facts may be established with respect to the choices made by an *S* at a choice point, the observed relative frequency of choices of a given kind on the one side, and the order in which the various alternatives make their appearance on the other. An important fact about the study of problems concerned with the *order* in which choices are made is that questions of fact about *order* cannot be answered by a study of observed relative frequencies alone. Such questions depend for their answer upon the comparison of the observed frequency in which a given sequence occurs with a mathematical expectation calculated on the basis of an acceptable hypothesis. That this is so follows from these considerations. Consider an urn containing $N(p + q)$ balls, p and q giving the proportion of white and black respectively. Let p and q be constant throughout the course of the drawings. Then the probability of sequences white-white is p^2 and the expected number of such sequences will be $(N - 1)p^2$. Any departure from this expectation leads to an increasing suspicion of bias in the drawings, as the departure increases. But $(N - 1)p^2$ varies with p. Hence, no observed quantity of sequences white-white can be taken either as indicative of avoidance of repetition or of stereotyping of white. What must be shown is an expected number of sequences with which the observed number can be compared. The calculation of such an expectation in the illustration is quite simple if the assumptions implied in the schema are realized (1). Since the assumption that p and q are constant is of such great importance, it is equally important that it not be made unless it can be submitted to statistical test. In most cases it can be put to test by Lexis analysis (11).

The requirement that observed quantities be compared with expected quantities has been variously met in the recent experimental

literature of the subject. Arons and Irwin (1) with psychophysical judgments calculated exact expectations. Krechevsky (8) with mazes compared his observed frequencies with expectations calculated on the basis of *a priori* hypotheses although the nature of his hypotheses in some instances is not clearly stated. Heathers (6) observed that under no conditions of choice (granted constant conditions of preference for a given alternative) could alternation by chance exceed 50 percent of the turns taken and used an arbitrarily selected 50 percent as his *a priori* expectation. Bousfield and Barnes (3) compared their series of 10 choices with distributions obtained by tossing sets of 10 coins as well as with the distribution obtained by the expansion of the binomial.

In the event that under the same experimental conditions the individual S maintains constant his proportion of choices of a given alternative it is easy to prove that taking all returns from the N Ss as a single set, all sequences comprising the same assortment of alternatives may be regarded as equally likely to appear, no matter how the Ss vary among themselves in respect of their observed relative frequencies of choices of a given kind. Consider M urns containing white and black. Let p_i be the proportion of white in the ith urn. Let p vary among the urns such that no two p's are necessarily alike. Let $q_i = 1 - p_i$. For the ith urn the probability of r whites and $(n - r)$ blacks (n balls being drawn) is clearly $\frac{n!}{r!(n-r)!} p_i^r q_i^{(n-r)}$. But in the event that the drawings from the various urns individually meet the Bernouilli requirement, this latter number may be written as the sum of $\frac{n!}{r!(n-r)!}$ values of $p_i^r q_i^{(n-r)}$, each of the latter elements being the likelihood of the appearance of one of the $\frac{n!}{r!(n-r)!}$ ways in which r whites and $(n - r)$ blacks can be coupled. Since these $\frac{n!}{r!(n-r)!}$ ways are all equally likely, it follows that summing the probabilities corresponding to each of the orders composing a given assortment of r white and $(n - r)$ black over the series of M urns previously defined will produce a set of $\frac{n!}{r!(n-r)!}$ identical sums, each of the sums containing M items of the type $p_i^r q_i^{(n-r)}$.

From this fact it follows that variation in the proportion from urn to urn (or from S to S in the psychological experiment) will not affect the equal likelihood of sequences reflecting the same proportions of black and white but differing in respect of the order in which the whites and blacks appear. Thus, in the Zenith study reported by

Goodfellow (**5**) it is permissible to erect the hypothesis that all orders of the type 4 white, 1 black, including BWWWW, WBWWW, WWBWW, WWWBW and WWWWB (B = Black and W = White) are equally likely. However, it must be remembered that nothing can be said (on the grounds of *a priori* probability) in respect of the relative frequency of 5 whites as against 4 whites, 1 black, since already we have shown that unless p_i is constant the distribution of the various assortments of white and black (disregarding order) is not Bernouillian in character and expectations cannot be calculated on that hypothesis. However the demonstration that all members of the set r white and $(n - r)$ blacks are equally likely, even under Lexis conditions over the M urns, is important because it permits a direct application of a test of statistical homogeneity in the absence of the theoretical expectation of the assortment r white and $(n - r)$ black.

That characteristic of action at a choice point, pertaining to the order in which choices are made, which has attracted most attention has been the avoidance of repetition of response. Theories of the avoidance of repetition have been of a number of kinds. One set of theories supposes the phenomenon to arise mechanically from the consequences of the previous choice. Among these theories is the one propounded by Ballachey and Buel (**2**) and exploited by Witkin and Schnierla (**15**), which explains the avoidance of repetition of a previous choice in a maze on the grounds that the geographical position of the rat in the maze after a given turn is taken makes easier the taking of the alternative turn at the next choice point. A second theory of avoidance of repetition was proposed by Dodge (**4**) and has been used variously by Thorndike (**14**), Telford (**13**), and Heathers (**6**). This theory explains the phenomenon as a consequence either of refractory phase (**4**) or of a factor analogous to refractory phase (**6**). Supporting this theory is very little psychological fact, although recently Heathers has reported that the avoidance of repetition of maze turns tends to diminish as the time interval between choices is increased. In the making of psychophysical judgments the same fact was reported by one of the present authors (**10**).

A third theory has regarded the avoidance tendency as similar to an attitude in its functioning. In support of this theory are two facts. In the first place the avoidance tendencies in the psychophysical experiment are very complex in their organization. A given report affects not only the next report but the one following and the one after that (**9**). In the second place the avoidance tendencies appear to affect not a unique response but a class of responses. *Ss* who compare alternately pairs of weights and pairs of sounds tend to avoid reporting 'louder' on the sounds if they have reported 'heavier'

on the weights (7). What establishes the limits defining the general class of responses was not indicated by the results of this experiment, although it would be no dangerous assumption to make to suppose that the *Ss* were avoiding categories of response, namely, 'greater than,' 'equal to' and 'less than.'

Somewhat similar to the problem of avoidance of repetition is the problem studied separately by Goodfellow (5) and Yacorzynski (16). These latter investigators have seen the study of the order in which choices are made not as a matter of repetition or of stereotyping of the previous choice, but as a matter of preference for a series of choices as a whole. Goodfellow observed in his treatment of the Zenith telepathy data that the *Ss* gave a large number of sequences which he defined as asymmetrical, a fact which suggested to him that the *Ss* were attempting to make guesses which matched their naïve notions of probability theory. Thus, according to Goodfellow relatively few instances of RBRBR or of BRBRB occurred because the action of the *Ss* was regulated by a tacit assumption that a properly constructed roulette wheel would not frequently yield simple alternation of red and black. A theory of this kind, to put it generally, maintains that the action of the *S* is determined by an hypothesis (in this case irrelevant to the main purpose of the choice requirement). There is also involved the assumption that the *Ss* find their task (*i.e.*, engaging in extra-sensory perception) a meaningful activity.

If the theories just enumerated be examined comparatively, it appears that only the second, third and fourth can make any claims to generality; the theory of centrifugal swing is restricted in its application to maze running. The attitude and the hypothesis theories have in common the fact that they regard the phenomenon in question understandable only in the light of the action of the individual as a whole. What difference exists between the two theories is relatively minor at the present time, the one supposing that the *S* is oriented towards the individual choice, the other supposing that the *S* is oriented towards sets of choices. Contrasting with the theory that choice point action of this kind is understandable only in the light of the *S* as an organic unit is the third theory enumerated, which finds in the operation of the law of refractory phase the explanation for avoidance of repetition of response.

While the present experiment is not crucial for a choice among these various theories (if a choice is possible) it has been designed to collect more facts upon which a choice ultimately can be made.

Design of Experiment

In the present experiment the following requirements were established by considerations of psychological and probability theory:

1. A relatively large number of *S*s must each give a reasonably large number of series of choices as between 2 alternatives, in each case a series being composed of 5 choices.

2. The choices made by each *S* should be made under constant conditions for any one series but the conditions should be varied from series to series in order to permit the selection of invariant features of the behavior observed over the entire set of series.

3. Sources of experimental variation should be introduced in such a manner as to cast a significant light upon the factors affecting the tendency to depart from chance frequencies.

To satisfy these requirements the experiment was cast according to the principles of factorial design. While certain of the data yielded were susceptible of treatment by the analysis of variance most of them were not. This fact however has little bearing upon the choice of factorial design, since the advantages of that type of design are not confined to the fact that such a design makes possible the analysis of a variance.

To require the *S* to make 5 choices in succession a maze plan was devised with the following properties: (a) any *S* traversing the maze would arrive successively at 5 choice points; (b) at each choice point one of two directions was possible, either right or left (retracing was prevented by instruction); (c) all turns involved in the maze except those at the choice points (which required a 90° turn) were plotted as a curve with radius the maximum possible for the size of maze used; (d) the distance from choice to choice point was constant, being $3\frac{3}{4}$ inches in one maze and $7\frac{1}{2}$ inches in a second.

In order to encourage a maximum of variability in the kind of pattern exhibited in the successive runs and to afford some insight into the factors affecting the results the following experimental variations were employed: (a) Two sizes of maze were constructed, the distance from choice point to choice point in the one maze being twice the corresponding distance in the other.[1] This variation in experimental condition affected the time elapsing between choices by doubling it (approximately) in the longer maze as against the shorter. (b) To effect a change in the preference shown (as between right and left turn) a spatial vector was used. This was provided by the construction of a tilted board with recess to carry the maze. The board (and thus the maze) could be tilted either to left or right at an angle of 45°. (c) To introduce a directional vector which relied upon suggestion rather than upon any influence exerting a direct effect upon the performance, the *S* was subjected in certain runs to one of two noise conditions. Under the one condition a 1000 cycle tone at a relatively high intensity level was directed into the left ear, under the other a tone of the same intensity was directed into the right ear.[2] (d) There being two lengths of maze, three noise conditions (including silence) and three tilt conditions (right, left and flat) it is clear that 18 combinations exhaust the ways in which one of each condition could be represented in the experimental conditions confronting the *S* on a single run. These 18 conditions are listed together with symbols identifying them in Table 1.

It is apparent that these 18 runs can be ordered in 18! ways. So that the order in which they were done might be eliminated as a factor in the results, and to establish conditions favorable to the study of the effect of duration of the experiment upon the results, a table was prepared according to the following plan: (a) Each of the 18 runs was assigned a letter (as is the case in Table

[1] The maze pattern was drawn upon a piece of plywood. Awlholes were placed in the patterns at intervals of 1/4 in. Into the awlholes were driven escutcheon pins, each pin having a rounded head 3/32 in. in diameter. The pins were carefully aligned and driven to a uniform distance in the plywood base. The design was such that each choice point was marked by a pin. The *S* traced the maze by moving from pin to pin with the forefinger of his right hand.

[2] The tones were generated by an electrically driven tuning fork, the oscillations of which actuated a telephone receiver. The *S* wore a headset with two receivers only one of which contained a diaphragm, the other remaining dead. In order to control the ear into which the sound was directed, the receivers were interchanged. When a series was run in which the sound was not used, the *S* continued to wear the headset, the tuning fork remaining silent. The energy level fixed at the outset of the experiment was maintained constantly for all *S*s and was the same for both ears in the case of each *S*. Although it was measurable, the level was not determined since nothing in the experiment depended upon the outcome of the measurement. The level was selected on the basis of the report of the experimenters and interested colleagues who described it as loud, approximating a condition of uncomfortable loudness. Sponge rubber pads were cemented to the telephone receivers to increase the comfort of the *S* and to reduce the possibility of bone conduction in the ear not subject to auditory stimulation.

TABLE 1

COMBINATIONS OF EXPERIMENTAL CONDITIONS

Symbol	Combination
A	Short Maze Left Tilt Left Ear
B	Short Maze Left Tilt Right Ear
C	Short Maze Left Tilt Silence
D	Short Maze No Tilt Left Ear
E	Short Maze No Tilt Right Ear
F	Short Maze No Tilt Silence
G	Short Maze Right Tilt Left Ear
H	Short Maze Right Tilt Right Ear
I	Short Maze Right Tilt Silence
J	Long Maze Left Tilt Left Ear
K	Long Maze Left Tilt Right Ear
L	Long Maze Left Tilt Silence
M	Long Maze No Tilt Left Ear
N	Long Maze No Tilt Right Ear
O	Long Maze No Tilt Silence
P	Long Maze Right Tilt Left Ear
Q	Long Maze Right Tilt Right Ear
R	Long Maze Right Tilt Silence

1). (b) In the table each letter appeared in the same position the same number of times. Since there were 18 positions (one for each run) this multiple (established at 6) fixed the number of *Ss* to be used in the experiment, namely 108. The finished table consisted in 108 columns of 18 letters each. Use of the table resulted in the following consequences: (a) Each of the experimental conditions occurred in each of the 18 temporal positions exactly 6 times. (b) 108 *Ss* were required for the experiment. (c) Each *S* gave 18 runs, one under each of the 18 conditions listed in Table 1. (d) A total of 1944 runs were taken over the entire 108 *Ss*. (e) Since each run required 5 choices, a total of 9720 choices were made. (f) Since each condition appeared in first place 6 times, second place 6 times, etc., it was possible to establish as a fourth variable the duration of the experimental period preceding a given run.

The advantage of this experimental design lies principally in the fact that not only is a large amount of variation introduced into the conditions under which the *S* makes his various choices, but also in the fact that by appropriate classification of the runs the effect of length of maze, sound and spatial vectors and duration of the experimental period can be studied. It should be remembered that the nature of the design is such that data studied with respect to one set of experimental conditions are automatically exactly equated for all other conditions, *i.e.*, if we classify all runs with respect to length of maze, the two sets of data differed only in respect of length of maze, being taken under conditions equally affected by the noise and spatial vectors and being taken from runs equally represented at each of the 18 stages of the experimental hour.

MISCELLANEOUS ITEMS OF PROCEDURE

1. The *Ss* were blindfolded and great care was exercised that they never saw the experimental apparatus.

2. The instructions were as follows:

"I am going to ask you to run a series of mazes blindfolded. The mazes will differ and the conditions under which you will run the mazes will differ. At the conclusion of a number of practice or exploratory runs there will be a test run. You will not be told when the test run occurs; neither will you be told what the basis of scoring the test run is. Learn as much as you can about the construction of the mazes, subject however to the condition that you will be permitted to move only to the right, to the left, or forward. You will never be permitted to turn back. In tracing the maze, always use the forefinger of the right hand. Work carefully, not quickly and remember that each trial is a new experience. It is important that you bear in mind the fact that backward movements are forbidden."

3. After the 9th run a 5 minute rest period was given the *S*. During this interval the blindfold was removed after the *S* had been taken behind a screen shielding him from the apparatus.

4. The *Ss* consisted of 108 men and women, all students at the University of Pennsylvania. The experiment was executed at the Psychological Laboratory of the University of Pennsylvania from November to May of 1939–1940.

5. The order of presentation of runs assigned any *S* from the list of 108 orders was determined randomly.

Results

Analysis of Frequencies of Sequences RRRRR, RLRLR, LRLRL and LLLLL

Table 2 shows the frequencies with which right and left turns are observed at each of the 5 choice points, the data being organized such that at any of the choice points and in the instance of any run it may be seen by inspection which turns had been taken previously. Thus in the 1944 runs, 1016 turns were taken to the right and 928 to the left at the first choice point. Of the 1016 instances in which a right turn was taken at the first choice point, 679 turned right at the second choice point, the remaining 337 taking a left turn. In the table the values in row 5 give the frequencies with which the 32 sequences occurred. The 32 possible orders are assigned the numbers appearing in row 6 for purposes of identification. They may also be identified by examining the order of the turns apparent in the table.

Examination of Table 2 yields the following observations:

1. A striking tendency exists for sequences RRRRR, LLLLL, RLRLR and LRLRL (numbers 32, 1, 22 and 11) to occur. If the table be examined closely it is clear that at every choice point in the four series of choice points leading to RRRRR, LLLLL, LRLRL and RLRLR, the preponderance of choices is made in the appropriate direction. Thus of the 337 who chose RL, 201 chose RLR, of these 201, 122 chose RLRLR and of these 122, 68 chose RLRLR.

With respect to sequences RLRLR and LRLRL it has already been demonstrated that in the event chance alone was affecting their frequency they should be homogeneous in frequency with all others of the same composition in L and R. Examination of the homogeneity of the set 3R2L (consisting of 10 patterns each of which contains 3 Rights and 2 Lefts) indicates that chi-square is 55.7 with a *P* of less than .01. The set 3L2R yields a chi-square of 116.7 with a *P* of less than .01. It is apparent that these two sets 3R2L and 3L2R are not homogeneous and in view of the fact that RLRLR and LRLRL occur consistently in large excess of their expectation (on the hypothesis of statistical homogeneity), we may conclude that the tendency to give alternating sequences is reliably in excess of what would be expected on a chance hypothesis.

The large number of instances of RRRRR and LLLLL cannot be tested (as has already been mentioned) on the basis of *a priori* probability. Three experimental hypotheses must be distinguished.

TABLE 2

FREQUENCIES OF RIGHT AND LEFT TURNS TAKEN AT EACH CHOICE POINT, EACH FREQUENCY REFERRED TO PREVIOUS CHOICES MADE—1944 RUNS

C. P. No.	Turn																													
1	L																R													
	928																1016													
2	L								R								L								R					
	621								307								337								679					
3	L				R				L				R				L				R				L				R	
	513				108				197				110				136				201				159				520	
4	L		R		L		R		L		R		L		R		L		R		L		R		L		R		L	R
	385		128		61		47		75		122		47		63		78		58		122		79		66		93		131	389
5	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R
	334	51	76	52	30	31	19	28	40	35	91	31	21	26	17	46	53	25	25	33	54	68	37	42	31	35	45	48	58	73	56	333
6	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32

On the one side, the possibility must be considered that the preponderance of Sequences 1 and 32 reflects a tendency for our *Ss* to divide into types, the types being defined by a pronounced tendency to prefer the one alternative as against the other. In the event this hypothesis were verified, we would conclude that the tendency to give Sequences 1 and 32 was dependent upon the fact that we had selected *Ss* who fell naturally into two classes, those who strongly preferred right turns as such and those who strongly preferred left turns as such. A second hypothesis would suggest that these results are a direct consequence of certain of the experimental conditions, chiefly the tilting of the maze. A third hypothesis would suggest that the large numbers of sequences 1 and 32 observed reflect a tendency for the *Ss* to stereotype their choices, *i.e.*, to repeat the previous choice no matter what its character.

Table 3 gives the frequencies with which the 108 *Ss* yielded right turns over the set of 18 runs (comprising 90 choices). It is apparent

TABLE 3

DISTRIBUTION OF RIGHT TURNS TAKEN BY THE *Ss*

(Maximum 90, Minimum 0)

Number of Right Turns	Observed Frequency
74–81	2
66–73	8
58–65	15
50–57	16
42–49	30
34–41	13
26–33	6
18–25	12
10–17	4
2–9	2
	$N = 108$

from the table that over the whole set of 90 choices few *Ss* are observed to give either a large or a small number of right turns. A common assumption made in respect of data of this kind is that they should distribute according to the normal curve. Fitting a normal curve to the distribution in Table 3 yields a chi-square of 17.64 with P equal to .04 for 9 degrees of freedom. While we may look with suspicion at the normality of the distribution on the basis of this analysis, we may not with certainty reject the null hypothesis. The suspicion that the distribution may not be normal is due largely to the conditions obtaining at intervals 18–25 and 26–33, in the case of one of which a mode appears, while in the case of the other there is an abnormally low frequency. But these departures from the normal expectation do not justify the conclusion that the data are drawn from *Ss* representing types in respect of strong preferences for the

one or the other turn. Since the abnormal number of instances of 5R and 5L appearing in Table 2 cannot be attributed to the fact that the *S*s represent two contrasting classes with respect to constant preferences, its origin must be sought elsewhere.

In view of the fact that the foregoing analysis suggests the absence of types among the *S*s, we may raise the question as to whether the preponderance of sequences 1 and 32 is due to the tilting of the maze. The crucial problem in the handling of this question is the development of a sampling expectation. As shown heretofore it is not legitimate to assume a Bernouilli distribution of assortments in R and L (*i.e.*, the distribution to be predicted from the expansion $[p_r + p_l]^5$), principally because such a distribution of the assortments does not permit the variability existing among the *S*s in respect of tendency to choose *R* (evident in Table 3) to make its appearance. The form of the distribution appearing in Table 3 suggests, however, that with respect to any sub-sample of the data having the characteristics of being taken under constant experimental conditions a roughly normal distribution of right turns should be observed. Strictly, the present data cannot meet the requirements of furnishing sets of 108 runs taken under exactly the same conditions since the conditions of the experiment were confounded with temporal order, *i.e.*, each of the 18 conditions appeared in each of 18 temporal positions 6 times. However, Lexis analysis of the percentage of right turns taken from trial to trial gives a Lexis Ratio of 1.38, indicating that no grounds exist for believing that temporal order produces any excessive variability in the data in respect of tendency to turn to the right. Hence, we may safely consider all data taken under the same experimental conditions as homogeneous in respect of the conditions, despite the fact that in any set of 108 runs (one from each of 108 *S*s) taken under any one of conditions A to R as defined in Table 1, the runs existed in sub-sets of 6, one subset each being taken in each of the 18 temporal positions. Table 4 gives the frequencies with which the 6 assortments, 5R, 4R1L, 3R2L, 2R3L 1R4L and 5L occur under each of the 18 experimental conditions. The conditions are enumerated by letter and may be identified by reference to Table 1. The last two columns give chi-square and *P* for the hypothesis that the distributions fit normal distributions with Mean and S.D. equal to that observed in each case in the data. The hypothesis tested here is that subsets of the data should take the form of the data as a whole as shown in Table 3. It is clear from inspection of Table 4 that the departures from expectation are sufficiently large in most cases to warrant the rejection of the sampling hypothesis.

Examination of Table 4 makes it clear that the preponderance of sequences 1 and 32 is not a direct consequence of the experimental

TABLE 4

FREQUENCY OF ASSORTMENTS UNDER EACH OF THE 18 VARIATIONS OF EXPERIMENTAL CONDITIONS

Exp. Var.	5R	4R1L	3R2L	2R3L	1R4L	5L	χ^2	*P*
A	19	14	14	21	11	29	14.95	.01
B	13	10	15	21	19	30	10.85	.01
C	22	7	15	16	19	29	20.36	.01
D	22	15	24	17	12	18	8.35	.01
E	20	17	23	22	13	13	4.80	.05
F	26	17	16	20	10	19	11.57	.01
G	32	14	19	19	7	17	16.68	.01
H	21	29	19	16	8	15	5.94	.02
I	35	7	31	10	13	12	18.59	.01
J	13	9	24	20	19	23	5.79	.02
K	8	11	21	28	18	22	5.84	.02
L	12	11	15	23	17	30	11.82	.01
M	13	11	22	30	20	12	3.26	.10
N	8	11	27	31	18	13	1.93	.10
O	12	17	26	27	14	12	3.10	.10
P	18	22	23	23	12	10	2.85	.10
Q	20	21	21	21	10	15	6.24	.02
R	19	22	23	19	10	15	6.02	.02
Totals	333	265	378	384	250	334		

conditions, in particular the maze tilt. Conditions A, B, C, J, K, and L all have in common the fact that they included a left tilt to the maze. Conditions G, H, I, P, Q, and R all have in common the fact that they included a right tilt to the maze. By inspection it is apparent that modes occur both at 5R and 5L in A, B, C, G, and J, and that in every instance the frequencies in the least favored direction are quite large.

In conclusion of our consideration of Tables 2, 3, and 4, it is apparent that a significantly large number of sequences RRRRR, LLLLLRLRLR, and LRLRL occur, and furthermore that we must reject the hypothesis that RRRRR and LLLLL occur in large numbers because of the existence of types among our Ss or that they occur because of the bias introduced by the tilting of the maze. We are left with the conclusion that the excessive number of sequences 5R and 5L is due to stereotyping of the initially chosen response. We will entertain the hypothesis tentatively that the excessive number of sequences RLRLR and LRLRL is due to the avoidance of repetition.

Effect of Experimental Variation upon Frequencies of Critical Sequences

Table 5 gives the frequencies with which 22 of the 32 sequences are observed when the 1944 runs are fractionated according to the experimental variation. The 22 sequences include the 2 sets of 5 Rights and 5 Lefts respectively, and the 2 sets made up respectively of the 10 combinations of 3 Rights and 2 Lefts, and 3 Lefts and 2

TABLE 5

FREQUENCIES OF SEQUENCES COMPOSING ASSORTMENTS 5L, 3L2R, 3R2L AND 5R, FRACTIONATED ACCORDING TO EXPERIMENTAL VARIATION

Sequence	1 Trials 1–6	2 7–12	3 13–18	4 Left	5 Flat	6 Right	7 Long	8 Short	9 Right	10 Silence	11 Left	12 All Runs
LLLLL	99	122	113	163	87	84	152	182	108	117	109	334
LLLRR	20	15	17	21	15	16	36	16	21	12	19	52
LLRLR	14	9	8	13	10	8	22	9	8	13	10	31
LLRRL	6	7	6	9	9	1	11	8	8	4	7	19
LRLLR	15	9	11	8	16	11	23	12	12	11	12	35
LRLRL	28	34	29	33	39	19	37	54	32	30	29	91
LRRLL	3	7	11	8	6	7	9	12	7	6	8	21
RLLLR	15	6	4	4	11	10	16	9	12	6	7	25
RLLRL	7	8	10	4	12	9	20	5	7	10	8	25
RLRLL	21	17	16	16	21	17	33	21	19	16	19	54
RRLLL	10	9	12	13	8	10	15	16	13	7	11	31
RRRLL	16	22	20	20	14	24	31	27	17	23	18	58
RRLRL	14	16	15	11	11	23	12	33	15	16	14	45
RRLLR	11	10	14	8	11	16	16	19	15	15	5	35
RLRRL	15	13	9	8	14	15	19	18	13	11	13	37
RLRLR	17	26	25	14	28	26	45	23	21	22	25	68
RLLRR	12	12	9	9	15	9	23	10	11	8	14	33
LRRRL	3	4	10	5	6	6	7	10	4	5	8	17
LRRLR	10	5	11	10	14	2	15	11	12	4	10	26
LRLRR	13	8	10	12	12	7	16	15	10	13	8	31
LLRRR	7	11	10	7	13	8	18	10	8	9	11	28
RRRRR	102	117	114	87	101	145	123	210	90	126	117	333
%LRLRL	20.1	28.1	23.4	25.6	26.6	17.6	16.7	33.4	23.0	26.2	22.4	23.7
%RLRLR	14.4	20.3	18.8	13.5	20.3	19.1	22.3	13.1	16.7	17.5	19.8	20.4

Rights. Columns 2, 3, and 4 divide the data according to the trial in which the run occurred; columns 5, 6, and 7 divide the data according to the manner in which the maze was tilted; columns 8 and 9 divide the data according to the distance from choice point to choice point and columns 10, 11, and 12 divide the data according to the noise conditions obtaining. The final column gives the frequencies observed over the entire 1944 runs. In addition to the frequencies with which the various sequences are observed, the final two rows give (1) the percentage of the total frequency of 3 Rights and 2 Lefts represented by sequence RLRLR, and (2) the percentage of the total frequency of 2 Rights and 3 Lefts represented by sequence LRLRL. A principle of the table is the fact that all frequencies of a sequence in the case of a given fractionation must equal the value for that sequence in the final column.

Examination of Table 5 discloses the following facts:

1. Without exception both RLRLR and LRLRL occur in excess of 10 percent of the frequency of their respective classes (the 10

percent being a theoretical expectation), under all conditions of fractionation.

2. Without exception LRLRL occurs with the highest frequency (among members of its class) under all conditions of fractionation.

3. RLRLR occurs with the greatest frequency under all fractions of duration, on the flat and right tilted mazes, but not on the left-tilted maze, on the long but not on the short maze, and under both noise conditions but not under silence.

4. The effect of duration (fractionation given in Columns 1, 2, and 3) appears to manifest itself in an increase in the 4 sequences RRRRR, LLLLL, RLRLR, and LRLRL.

5. The effect of the tilt upon RRRRR and LLLLL is difficult (for technical reasons) of evaluation. If the maze is tilted to the right LRLRL is considerably diminished in frequency. If the maze is tilted to the left, RLRLR is likewise diminished in frequency.

6. The effect of shortening the maze is definitely to increase the percentage of LRLRL. The percentage of RLRLR falls off when the maze is shortened. Shortening the maze increases the number of RRRRR greatly and the number of LLLLL to a lesser extent.

7. The noise conditions appeared to have little effect upon any of the 4 critical sequences.

Discussion

The foregoing results may be discussed in the light of the theories enumerated earlier and in the light of results from similar experiments reported by Goodfellow (**5**), Yacorzynski (**16**), and Bousfield and Barnes (**3**).

Clearly the fact that the same experimental conditions produce both stereotyping and alternation of the response excludes a refractory phase hypothesis, since such an hypothesis, in explaining the excess number of alternations, would predict a decrease in the number of repetitions. For the same reason the peculiarities in the order in which the turns are taken in this maze are inexplicable on the grounds of the theories proposed by Witkin and Schnierla, Ballachey and Buel and others. We do not intend to maintain that centrifugal swing and allied forces will not result in excess in alternation; rather we maintain that the present case is inexplicable on the basis of such an hypothesis because it fails to take care of the excess number of both alternations and repetitions. In addition, the persistence of the repetition and the alternation in the face of the spatial vector (*i.e.*, the tilting of the maze) makes untenable hypotheses like centrifugal swing.

The present data exhibit a remarkable similarity to certain data reported by Yacorzynski in a pseudo-extrasensory perception ex-

periment. Using 20 manic-depressives and 20 schizophrenics, with 40 normals as controls, Yacorzynski duplicated the conditions of the usual extrasensory perception experiment, observing the frequencies in the orders of the reports of heads and tails, etc. In his report the 32 possible sequences to be observed when 5 successive guesses are made are collapsed into 16 (following Goodfellow's procedure), frequencies of sequences HHHHH and TTTTT being combined as a frequency of 11111, HTHTH, and THTHT being combined into 12121, etc. Under these conditions it is of course not possible to decide whether the frequencies of the 6 assortments (5H, 4H1T, 3H2T, 2H3T, 1H4T, and 5T) distribute in the manner observed for the analogous assortments in the present experiment. However, the theorem bearing upon the equal likelihood of the members of the individual assortments may be put to test with his data. It is apparent from the data that the experimental *S*s (the manic-depressives and the schizophreics) gave many more instances of 12121 (corresponding to our LRLRL and RLRLR) than the theorem allows. We also regard it likely from Yacorzynski's data that the number of repetitions exceed what should be expected under the hypothesis of a normal distribution of the assortments, although since the various sequences have been combined in the manner described above, it is impossible to put this latter question to statistical test. What is true of Yacorzynski's *control* subjects is also true of Goodfellow's Zenith Radio subjects, *i.e.*, fewer alternations are observed than are expected and the observed quantities of repetitions are very low. Again we find ourselves unable to evaluate the observed quantities of repetitions (with Yacorzynski's controls or Goodfellow's *S*s) because of the manner in which the data appear.

Similar in some respects to the outcomes of Yacorzynski's experiment and our own is the experiment reported by Bousfield and Barnes, in which the experimenters induced the *S*s to make nonsense choices between two alternatives. Comparison of their frequencies of 10 choices of a kind (similar to our case of RRRRR and LLLLL) indicates that for both varieties of choice a larger frequency is observed than is given by the expansion of the binomial. The evaluation of this fact is made difficult by reason of the fact that in the comparison a Bernouilli distribution of the relative number of choices is assumed in the use of the expansion of the binomial. However the pronounced modes at 0 and 10 frequency of the one alternative suggest very strongly the presence of true stereotyping of response. Furthermore, a large excess of the case 5 of the one alternative and 5 of the other appears in the distribution of frequencies of each of the 11 assortments, a fact suggesting (but by no means proving) the increase in frequency of simple alternation.

A consideration of the results from the four experiments suggests the hypothesis that the *S*'s orientation towards the task is an important consideration in the production of alternation and repetition. Affected by or concerned with this orientation are such matters as (1) whether the *S* finds the task meaningful or meaningless. This might also be stated as the extent to which the *S* accepts as real the goal involved in performance on the task. (2) the degree of satiation in which the *S* exists and (3) the demand made upon the *S* by the task, *i.e.*, whether decisions can be made subject to the *S*'s control or at a pace which is faster than is suitable to the *S*. According to this hypothesis *S*s engaged in making choices will be guided by hypotheses, as argued by Krechevsky and by Goodfellow, whenever the *S* finds the task meaningful. Whenever, however, the goal represented by performance on the task is rejected, the *S* will go over to a simple mechanical response, either repetition or alternation of the possibilities. Under such an hypothesis, satiation (loss of interest in the goal) and the increase in the difficulty of the task by increase of pressure upon the *S* (increase of number of choices per unit time) should have as a consequence an increase in alternation or repetition. This of course happens in the present experiment. In a general sense it would appear that avoidance of repetition and stereotyping are a species of position habit.

There remains for consideration a problem recently discussed by Skinner (**12**) and mentioned briefly in the present introduction, *i.e.*, the question whether *S*s in experiments requiring choices in sequence react to the sequence as a whole or to the individual components of the sequence. We do not regard this as a question involving a dichotomous choice, as does apparently Skinner, who prefers to regard Goodfellow's Zenith *S*s as operating at the level of the individual choice, rather than as avoiding the use of symmetrical sequences as Goodfellow called them. That *S*s may be avoiding repetition at the level of the single choice and also at the level of the sequence is a *logical* possibility since it can be shown by mathematical analysis that over a large number of *S*s the sequences LRLRL and RLRLR in the *n*th trial occur more often than any other member of the assortments to which they refer, and in addition within any given *S*, do not follow each other as frequently as their frequency demands in random sampling. It might be thought that the present data afford an opportunity of testing this merely logical possibility but such is not the case, since each of our 108 *S*s gave only 18 sets of choices, the choices being taken under varying conditions. For the calculation of the expected number of cases where the same *S* gives identical sequences on two successive runs more than 18 runs are needed and the conditions should be constant for all runs. However,

the logical possibility suggests the complexity of the processes involved in the guessing as between two alternatives and shows clearly the fact that Goodfellow's hypotheses cannot be rejected on the grounds which Skinner assigns.

(Manuscript received October 28, 1942)

References

1. Arons, L., & Irwin, F. W. Equal weights and psychophysical judgments. *J. exp. Psychol.*, 1932, **15**, 733–751.
2. Ballachey, E. L., & Buel, J. Centrifugal swing as a determinant of choice point behavior in the maze running of the rat. *J. comp. Psychol.*, 1934, **17**, 201–223.
3. Bousfield, W. A., & Barnes, J. L. Empirical probability in a free two-choice situation. *J. Psychol.*, 1941, **12**, 45–55.
4. Dodge, R. A note on Professor Thorndike's experiment. *Psychol. Rev.*, 1927, **34**, 237–240.
5. Goodfellow, L. D. A psychological interpretation of the results of the Zenith radio experiments in telepathy. *J. exp. Psychol.*, 1938, **23**. 601–632.
6. Heathers, G. L. The avoidance of repetition of a maze reaction in the rat as a function of the time interval between trials. *J. Psychol.*, 1940, **10**, 359–380.
7. Irwin, F. W., & Preston, M. G. Avoidance of repetition of judgment across sense modalities. *J. exp. Psychol.*, 1937, **21**, 511–520.
8. Krechevsky, I. "Hypotheses" versus "chance" in the pre-solution period in sensory discrimination-learning. *Univ. Calif. Publ. Psychol.*, 1932, **6**, 27–44.
9. Preston, M. G. Contrast effects and the psychophysical judgments. *Amer. J. Psychol.*, 1936, **48**, 389–402.
10. Preston, M. G. The effect of interpolated time intervals upon the contrast effects. *J. exp. Psychol.*, 1936, **19**, 706–715.
11. Reitz, H. L. *Mathematical Statistics*. Chicago: Open Court, Ill., 1929, Chapter VI.
12. Skinner, B. F. The processes involved in the repeated guessing of alternatives. *J. exp. Psychol.*, 1942, **30**, 495–503.
13. Telford, C. W. The refractory phase of voluntary and associative responses. *J. exp. Psychol.*, 1931, **14**, 1–36.
14. Thorndike, E. L. The refractory period in associative processes. *Psychol. Rev.*, 1927, **34**, 234–236.
15. Witkin, H. A., & Schnierla, T. C. Initial maze behavior as a function of maze design. *J. comp. Psychol.*, 1937, **23**, 275–304.
16. Yacorzynski, G. K. Perceptual principles involved in the disintegration of a configuration found in predicting the occurrence of patterns selected by chance. *J. exp. Psychol.*, 1941, **29**, 401–406.

STUDIES IN DECISION: I. DECISION-TIME, RELATIVE FREQUENCY OF JUDGMENT AND SUBJECTIVE CONFIDENCE AS RELATED TO PHYSICAL STIMULUS DIFFERENCE *

BY LEON FESTINGER

State University of Iowa

A quantitative theory of decision has been presented elsewhere (3) with a precise analysis for decisions in which a choice must be made between two possible alternatives. This analysis should hold for any such situation. To test this theory empirically a psychophysical experiment was used for the following reasons: (1) In the field of psychophysics, we are able to set up situations which we can more adequately control. (2) We are able, in this area, to obtain quantitative measurements to a high degree of precision.

Precise predictions were made as to the functional relationships between decision-time and relative frequency of judgment as dependent variables and subjective confidence of judgment as the independent variable. The empirical evidence relevant to these precise predictions is reported in another article (4). The present article deals with evidence relevant to grosser predictions where the independent variable is size of physical stimulus differences. It also deals with the relation between size of stimulus difference and subjective confidence of judgment.

Procedure

A method of constant stimuli was employed in a two-category experiment in which a judgment was to be made as to whether one line was longer or shorter than another line.

Apparatus

Two black metal boxes were used for presentation of the stimuli. Ordinary electric light bulbs were placed inside each box. In front of each box was an opening, $\frac{1}{8}$ in. wide, extending the length of the box. A metal strip, placed through grooves on each side of the opening, could be moved back and forth so as to cover up all or any desired portion of the open part. The opening itself was covered with heavy ground glass through which the light could shine.

The boxes were mounted upon a table 31 in. high so that the openings in the front of the boxes were vertical. The distance between the neighboring edges of the two vertical lines was 9 in. A black cardboard screen with two slits in it was mounted in front of the apparatus so that

* This constitutes part of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Child Welfare, in the Graduate College of the State University of Iowa.

The writer is indebted to Dr. Kurt Lewin and Dr. Dorwin Cartwright for aid and suggestions.

the only thing visible from the front was a homogeneous black background and the two vertical lines. The experiments were always carried out in total darkness so that nothing could be seen until the vertical lines were illuminated.

The *S*s sat in a chair 130 in. away from the apparatus and midway between the two vertical stimuli.

The two lines were presented simultaneously to the *S* for 0.5 sec. The duration of the presentation of stimuli was controlled by an electron tube tachistoscopic timer. When the tachistoscope was turned on, thus illuminating the vertical lines, an electric stop clock started. After 0.5 sec. the lights went off automatically, but the clock kept running until the *S*'s response activated a voice key which stopped the clock. The electrical circuit used is shown in Fig. 1.

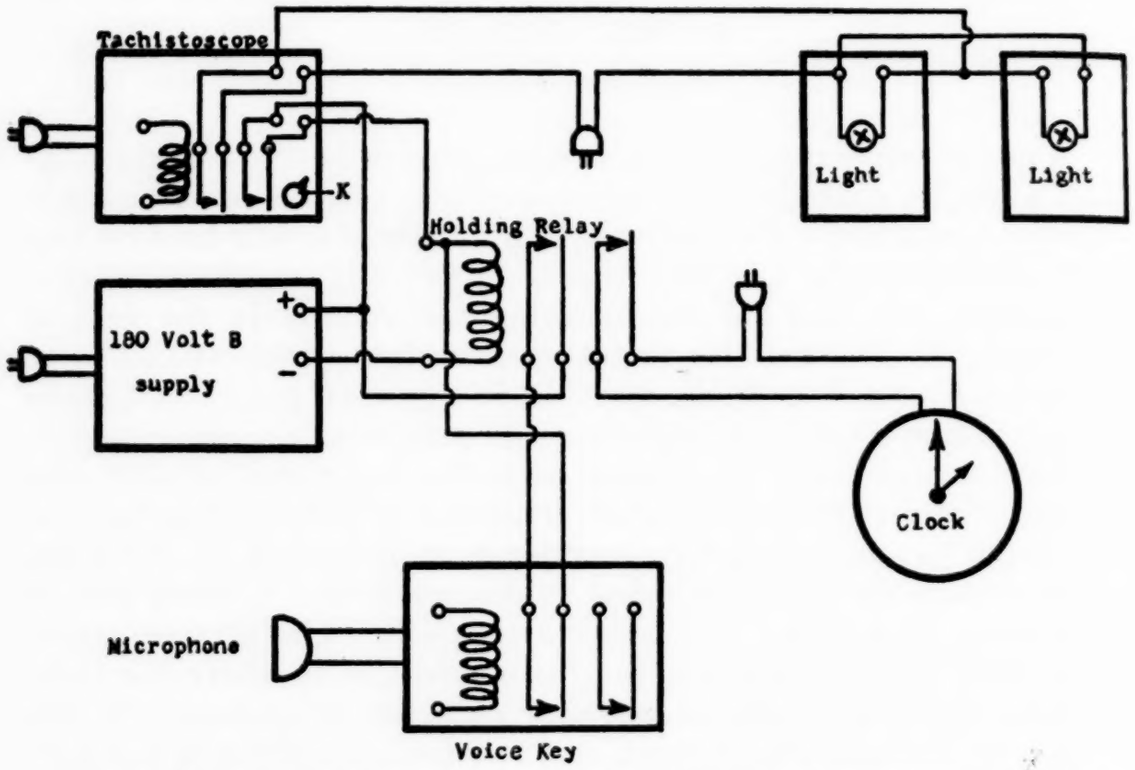

FIG. 1. Electrical circuit used for presentation of stimuli and measurement of decision-time. Sequence of operations: 1. Operate K to start Tachistoscope. This turns on lights and closes Holding Relay, thus starting clock. 2. Release K, resetting Tachistoscope, but Holding Relay remains locked. 3. Subject responds, Voice Key opens Holding Relay; clock stops.

After each response, the experimenter, using a very dim flashlight, recorded the proper data and set the apparatus for the next presentation.

The stimulus line at the left of the *S* was always kept at a length of 7½ in. This we shall refer to as the standard. The length of the other stimulus line was adjusted by means of the sliding metal strip which had a convenient matrix to its back. This line we shall call the variable stimulus.

Fifteen positions of the variable stimulus were used, seven positions shorter than the standard, seven positions longer than the standard, and one position equal to the standard. In sixteenths of an inch difference between the variable and the standard, those fifteen positions were: − 16, − 12, − 8, − 6, − 4, − 2, − 1, 0, + 1, + 2, + 4, + 6, + 8, + 12, + 16. Thus, the variable ranged from an inch longer than the standard to an inch shorter than the standard. This range is much larger than that used in the psycho-physical experiments, but was necessary here to make sure that the 'base value' for decision-time (immediate decision) was reached. That is, for purposes of fitting to our theoretical functions it is necessary to know the duration of an 'immediate' decision. It may also be noticed that the variable positions are concentrated mainly around the point of objective equality. This was felt to be advisable, since the theoretically predicted curves show the most rapid acceleration in this area.

The variable stimulus positions were presented in the following prearranged order: − 6, − 2, − 12, + 2, + 12, + 4, − 4, − 16, + 8, 0, − 8, − 1, + 1, + 16, + 6, − 8, + 12, 0, + 4, − 12, − 4, + 8, − 1, − 16, + 6, − 6, + 1, − 2, + 16, + 2. After this order was completed the reverse order was given and then the whole forward and reverse sequence continued.

Conduct of the Experiment

Five *S*s were used in the experiment. Each *S* made judgments under four different sets of instructions. For each of the four conditions, each *S* made 40 judgments at each variable stimulus position, making a total of 2400 judgments by each *S*. The experimental sessions rarely lasted over an hour for any subject because of the fatiguing nature of the experiment. An average of about 120 judgments could be made in one hour.

1. The condition which shall hereafter be referred to as the 'Usual Condition' was the first to be given. The instructions to the *S* for this condition follow:

Instructions

When the room is darkened, you will see two vertical lines illuminated for a period of time. You are to make a judgment as to whether the line on the right is longer or shorter than the line on the left. If the line on the right is shorter, you will say 'shorter'; if the line on the right is longer, you will say 'longer.' Do not say any other words before the judgment of longer or shorter. When you give your judgment, enunciate the word clearly and distinctly and fairly loudly.

After making each judgment you are to express your confidence that the judgment was correct in the following terms: 50 percent confidence means average or medium confidence, the amount of confidence one usually has in most of one's judgments; 75 percent means confidence definitely above average, the amount of confidence you have when you are quite sure you are right; similarly, 25 percent means definitely below average confidence, the amount you have when you are only guessing; 100 percent confidence is at the extreme end of the scale of confidence and means complete certainty, no possibility of error; similarly, zero confidence is at the other end of the scale of confidence and means no confidence in your judgment whatever; you are just as likely to be wrong as right.

Thus, you are to construct a 'mental scale' of confidence with 50 percent in the middle, 25 percent and 75 percent on each side of the middle range, and 100 percent and zero percent way out at the ends of the scale.

Sometimes subjects report that, in the process of making a judgment, their confidence fluctuates considerably over an appreciable range. This happens rarely but if it does happen, then the confidence rating which you give is to be approximately an average of the fluctuations which occurred.

The last two sentences in the first paragraph were inserted because of the use of a voice key to stop the clock. The second and third paragraphs, describing the confidence scale, are repeated verbatim from Johnson (5).

Before the actual experimentation began, the *S*s were given sufficient practice trials to accustom them to the situation. No arbitrary limit was set on the number of practice trials. They were continued until the *S*s were sufficiently at ease with the procedure.

2. The 'Accuracy Condition' followed upon completion of the 'Usual Condition.' The following additional instructions were given the subjects:

Accuracy Instructions

In this part of the experiment, you are to make your judgments as accurately as you can. See if you can be even more accurate than you were before. Try to make no errors. Remember, accuracy is the important thing here.

The above instructions are assumed to increase the restraining forces opposite to leaving the region of decision over what they were under the 'Usual Condition.'

3. The 'Speed Condition' was administered after the 'Accuracy Condition' with the following instructions:

Speed Instructions

In this part of the experiment you are to make your judgments as rapidly as you can. Remember, speed is the most important thing in this part of the experiment.

These instructions are assumed to lower the restraining forces below what they were in the 'Usual Condition.'

4. The 'Constant-Error Condition' came last. The instructions were as follows: (These instructions were presented as though specific for the subject involved. For three of the *Ss* they were precisely as stated below. For the other two subjects the word 'shorter' should be substituted for the word 'longer' wherever it occurs and *vice versa*.)

Instructions to .....

In all of the three preceding variations you have consistently been making an error in saying 'shorter' too many times. Even when the line was slightly longer you would very frequently say 'shorter.'

I would like you to try not to make this error again in this variation. Please bear in mind that you have a tendency to say 'shorter' which is producing this error in your judgments and please try to correct this tendency.

See if you can eliminate this error.

These instructions are assumed to increase the restraining force against making a decision in one direction more than in the other. The instructions as stated above should make the restraining force against "saying 'shorter'" greater than the restraining force against "saying 'longer.'"

Description of the Subjects

The five *Ss* used in the experiment were: Mr. Ben Willerman (B. W.), a graduate student in the Child Welfare Research Station; Miss Helen Sands (H. S.), a head teacher in the Iowa Preschool Laboratories; Miss Johanna Boet (J. B.), a graduate student in the Child Welfare Research Station; Miss Nancy Phillips (N. P.), a graduate student in the Department of Psychology; and Miss Grace Shattyn (G. S.), a graduate student in the Department of Psychology.

All of the *Ss*, with the exception of B. W., were completely naïve as to the purpose of the experiment and knew nothing of the theory involved. Subject B. W. was acquainted with the purpose of the experiment and knew the general outlines of the theory.

In the 'Constant-Error Condition' *Ss* B. W., H. S., and G. S. were given the instructions as stated above. The other two *Ss* were given the instructions with 'longer' and 'shorter' reversed.

Data Recorded

For each judgment three measures were recorded; namely, the judgment 'longer' or 'shorter,' the time from the start of the presentation of the stimuli to the response by the *S*, and the *S*'s rating of confidence.

Relation Between Confidence and Stimulus Differences

The data gathered in this experiment have been organized for an evaluation of the derivations made from the theory (3). Before presenting these results, however, we must digress momentarily for a consideration of the confidence judgments.

The relationship between confidence in one's judgment and stimulus differences is intimately connected with the derivations from the theory. These are actually derivations of the functional relationships holding between decision-time, relative frequency of judgment, and subjective probability (confidence) as to the correctness of the judgment. The derivations can be applied to stimulus differ-

ences instead of subjective probability only in so far as there is a monotonic relationship between the two. Furthermore, before a more precise test of the derivations can be made (**4**), this relationship must be exactly determined. Even for the gross derivations to be handled in the present paper it must be demonstrated that the relationship exists. We shall therefore proceed to do this.

Johnson (**5**) found that an ogive (the integral of the normal curve) best described the relationship between confidence and physical stimulus differences. This was found to be true for our data too. Figs. 2 through 6 show some of the ogives fitted to the empirical data of confidence. The examples shown are quite characteristic of all of the confidence functions. The fitting was done by the method of moments, without weightings. The means and standard deviations of the ogives are given with the figures.

Although these ogives are the 'best fitting' curves for the data, certain sections of the data show consistent deviations from the ogive. These deviations are similar for the same *S* in all four conditions.

Subject H. S. (Fig. 2) shows the following deviation: the empirical curve begins to rise slightly later than the ogive, then rises more sharply than the ogive, and then levels out very rapidly. Subject B. W. (Fig. 3) shows no deviation from the ogive pattern. All the other *S*s show a more rapid initial rise than the ogive, a slight leveling off at the point of subjective equality, and then again a rise more rapid than that of the ogive.

There are two ways of interpreting these deviations. One might argue that the ogive is not the correct functional relationship between the two variables concerned. Or, one might say that the ogive relationship is correct and that these deviations are due to the fact that the *S* used subjective units of different size in different parts of the confidence scale. If the true relationship between the confidence scale and stimulus differences were an ogive, then any non-linearity in the use of the confidence scale by the *S*s would result in deviations from the ogive relationship.

Subject H. S. used steps of 10 (*e.g.*, 70, 60, 80, etc.) in the confidence scale between confidence of 50 and 100. Below 50, however, the step suddenly increases so that except for an occasional 10, the only values used are zero and 25. If the sizes of these steps can be taken as symptoms of the size of psychological units at the different parts of the scale, such a distortion of the linearity of the scale would produce the type of deviations from the ogive observed in the case of this *S*.

Subject B. W., who shows no deviations from the ogive relationship, does not show any systematic differences in the unit of discrimination used on the confidence scale.

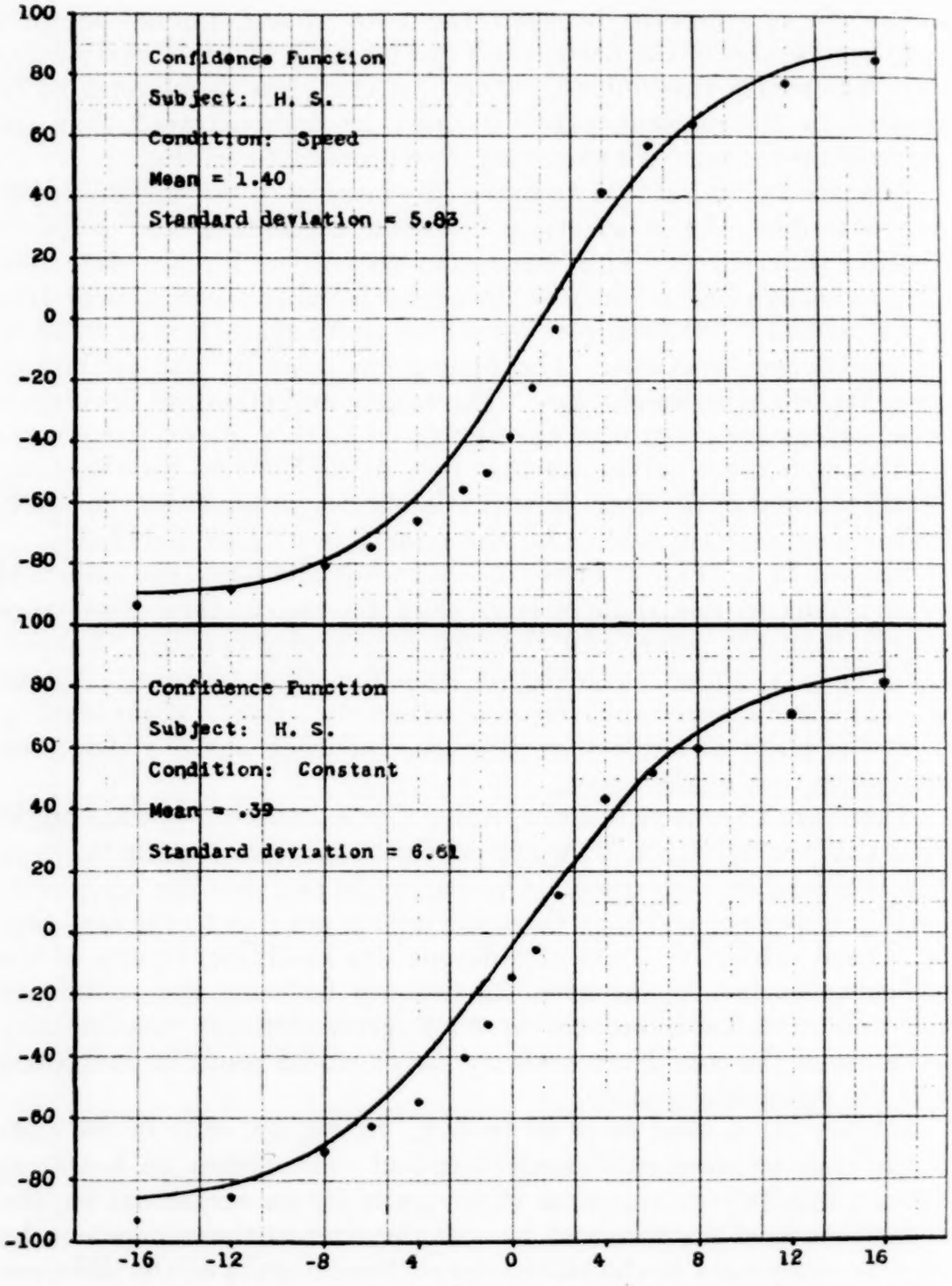

FIG. 2. (Top) Relation between confidence and physical stimulus-difference units for subject H. S. under 'speed' instructions. (Bottom) Relation between confidence and physical stimulus-difference units for subject H. S. under 'constant-error' instructions.

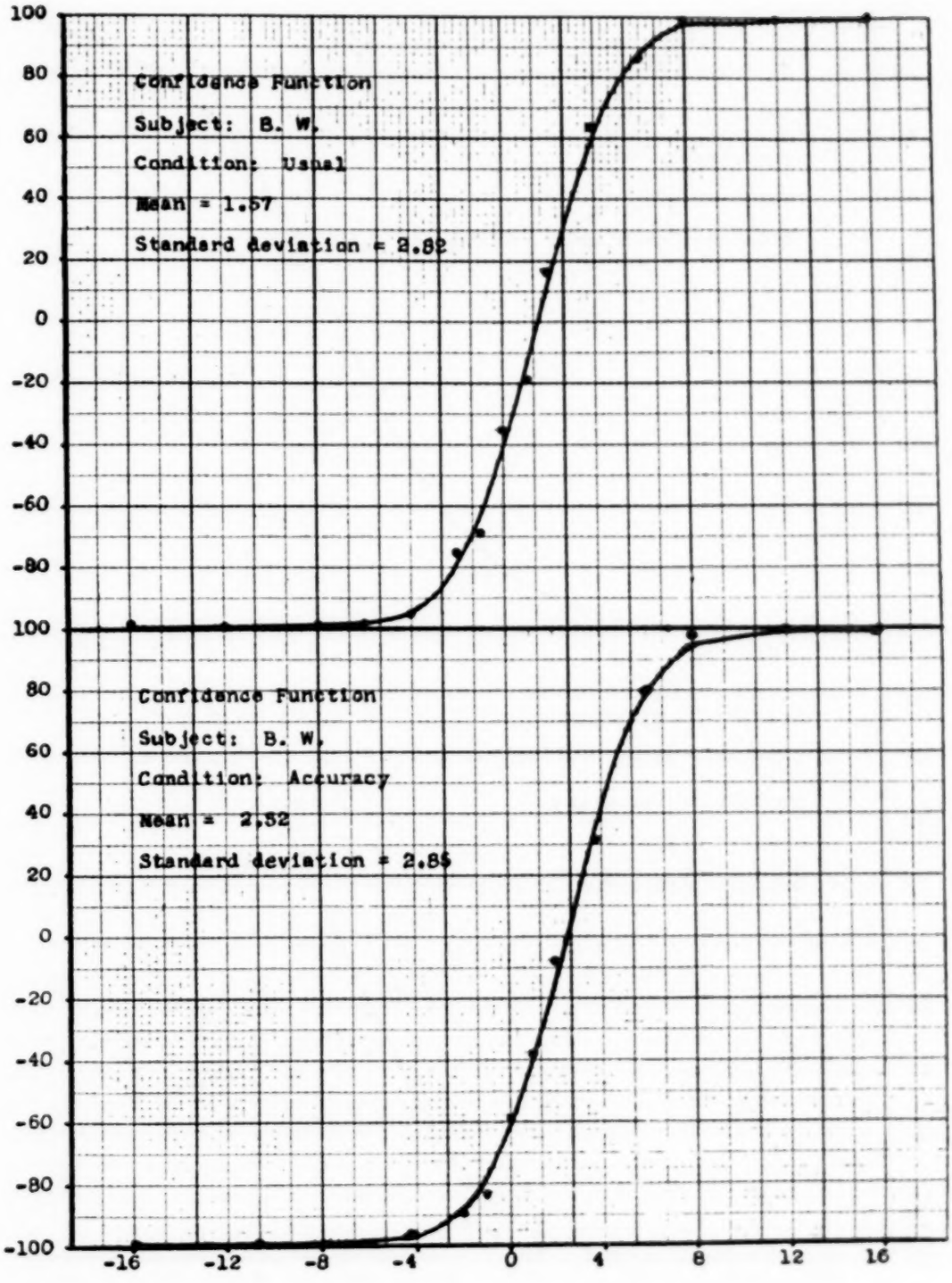

FIG. 3. (Top) Relation between confidence and physical stimulus-difference units for subject B. W. under 'usual' instructions. (Bottom) Relation between confidence and physical stimulus-difference units for subject B. W. under 'accuracy' instructions.

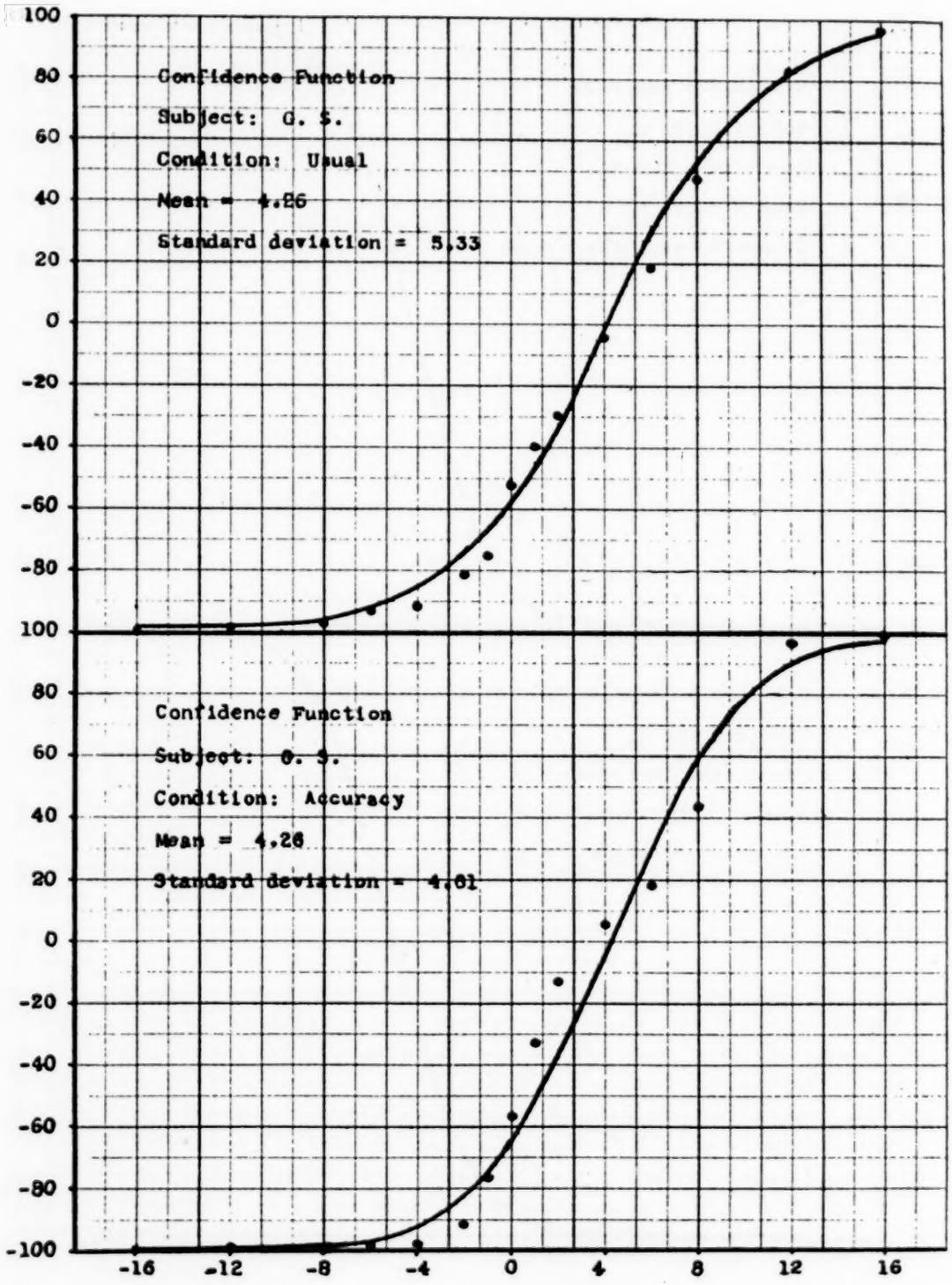

FIG. 4. (Top) Relation between confidence and physical stimulus-difference units for subject G. S. under 'usual' instructions. (Bottom) Relation between confidence and physical stimulus-difference units for subject G. S. under 'accuracy' instructions.

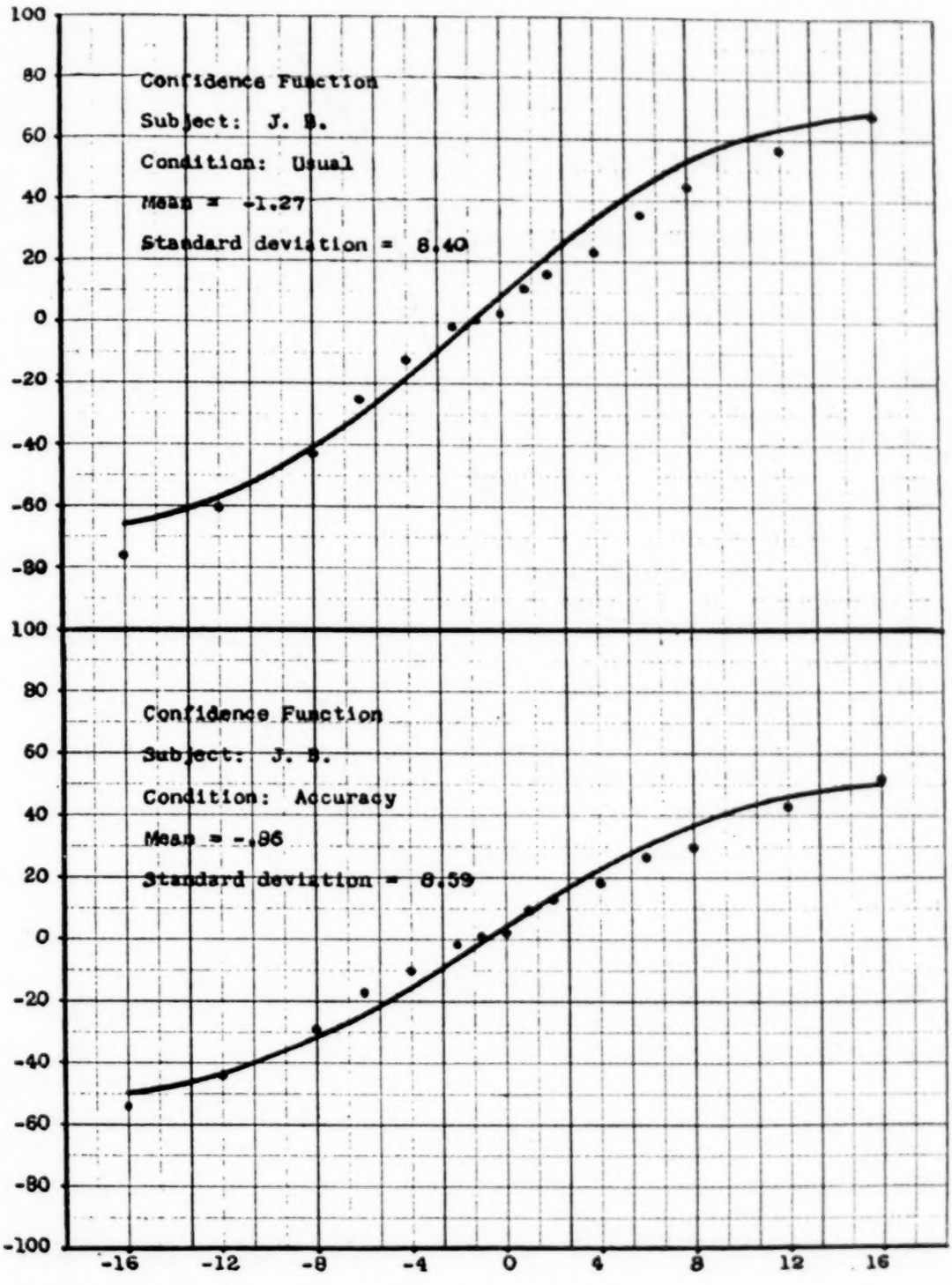

FIG. 5. (Top) Relation between confidence and physical stimulus-difference units for subject J. B. under 'usual' instructions. (Bottom) Relation between confidence and physical stimulus-difference units for subject J. B. under 'accuracy' instructions.

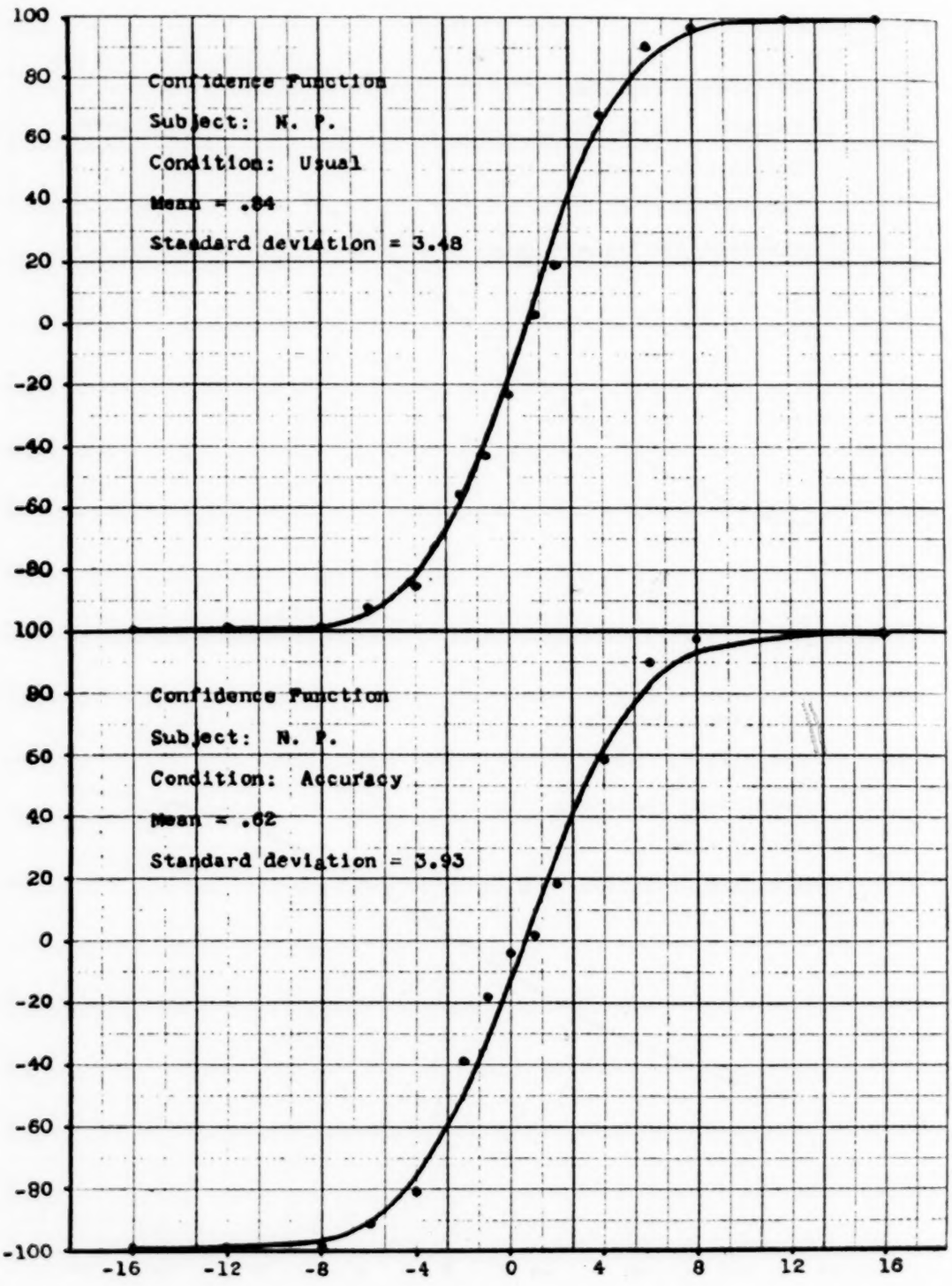

FIG. 6. (Top) Relation between confidence and physical stimulus-difference units for subject N. P. under 'usual' instructions. (Bottom) Relation between confidence and physical stimulus-difference units for subject N. P. under 'accuracy' instructions.

The other three *S*s, who show a deviational trend of the data from the fitted ogive which is the opposite of that shown by Subject H. S., all tend to show confidence discriminations which are finer near zero than near 100. Thus, subject G. S. usually uses steps of 25 near 100 percent confidence, but will discriminate even in steps of 1 and 2 near zero confidence. Subject J. B. gave ratings of 50 and 75 almost exclusively at the end of the scale (this *S* practically never gave a confidence rating over 75) but discriminated in units of 5 around zero confidence. Subject N. P. shows the same tendency, but not in as clear a fashion, to discriminate in smaller units near zero confidence.

Thus, there is good reason to believe that the deviations from the ogive are due to non-linearity in use of the confidence scale.

It may be noted (Table 1) that the standard deviations of the confidence function for the same *S* frequently change from condition

TABLE 1

STANDARD DEVIATIONS OF THE CONFIDENCE FUNCTION

Subject	Usual	Accuracy	Speed	Constant
H. S.	6.10	5.06	5.83	6.61
B. W.	2.82	2.85	3.04	3.75
G. S.	5.33	4.61	4.18	4.66
J. B.	8.40	8.59	7.68	8.69
N. P.	3.48	3.93	5.63	4.33
Average	5.23	5.01	5.27	5.61

to condition. In some instances these changes are small, in other instances they are appreciable. The last row in this table gives the average standard deviations for all five *S*s.

The question, of course, presents itself as to whether there is any consistency in these changes and whether any meaning can be attached to them.

It is apparent from Table 4 that the changes from condition to condition are not consistent between *S*s. None of the differences between means are statistically significant as measured by the *t* test. It must then be concluded that the changes in the standard deviation of the confidence function are due to individual factors specific to each *S*, which for some undetermined reasons cause changes in their feelings of confidence.

We may conclude that the relationship between confidence and stimulus differences is monotonic; more precisely, that it describes an ogival curve. We may therefore proceed to extend our theoretical derivations to relate to stimulus differences, and examine our data with regard to the derivations.

Evaluation of Results

The theoretical derivations are repeated here together with the relevant results.

Derivation 1

As relative frequency approaches 50 percent, decision-time approaches a maximum.

Table 2 presents the necessary data for an evaluation of this statement. The distance in stimulus units (sixteenths of an inch) between

TABLE 2

Differences (in $\frac{1}{16}$ of an Inch) Between Maximum Decision-Time and 50 Percent Relative Frequency

Subject	Usual	Accuracy	Speed	Constant	Average
H. S.	1.42	.10 *	.17 *	1.17	.715
B. W.	.32 *	1.53	.82 *	.30 *	.743
G. S.	.42 *	.96 *	1.23	.57	.795
J. B.	1.00	1.44	.59	.10 *	.783
N. P.	1.08	.84	.72	.73	.843

* Maximum decision-time closer to 50 percent point than to any other stimulus point.

the highest value of decision-time and the 50 percent point of the relative frequency ogive is presented for each of the five *S*s, for all four conditions of judgment. The 50 percent point of the relative frequency ogive has been calculated using the Müller-Urban weights. The highest decision-time value should not be expected to correspond exactly with the 50 percent relative frequency point, since the decision-time value used always falls at one of the variable stimulus positions, while the 50 percent relative frequency point usually falls between two stimulus values. Therefore, it is a question of how many stimulus units lie between the two points. As can be seen from the table, in 8 of the 20 comparisons the maximum decision-time is at the stimulus position closest to the 50 percent point of relative frequency. With two exceptions, all the other maximum decision-times lie at the stimulus position which is next closest to the 50 percent point of relative frequency. The two exceptions are: subject H. S., 'Constant' condition; and subject N. P., 'Usual' condition. In both these instances the maximum decision-time is two stimulus positions away from the 50 percent point of relative frequency. In other words, in all but two cases the 50 percent point of relative frequency falls between two stimuli, one of which represents the maximum decision-time.

The data can also be evaluated from the point of view of the distance (in sixteenths of an inch) along the stimulus continuum between the two points. The last column in the table gives this aver-

age distance for each *S*. For no *S* is this average distance greater than one-sixteenth of an inch.

One can state that for either method of comparison, the correspondence between maximum decision-time and 50 percent relative frequency is quite close.[1]

Derivation 2

The decision-time curve begins to rise at a point farther removed from the 50 percent point than does the relative frequency curve.

This derivation is completely borne out by the data. In every case the decision-time keeps increasing while the relative frequency is still zero percent (or 100 percent). A discussion of this result, with complete data for each *S*, is presented in another article (4), where theoretical curves are fitted to the empirical data.

Derivation 3

The maximum decision-time increases with an increase in the magnitude of restraining force.

Table 3 presents the difference in seconds between the highest and lowest points of decision-time for each *S* in each of the four conditions.

TABLE 3

DIFFERENCE IN SECONDS BETWEEN HIGHEST AND LOWEST POINTS OF THE DECISION-TIME CURVE

Subject	Usual	Accuracy	Speed	Constant Error
H. S.	.64	.73	.56	.58
B. W.	1.85	2.46	1.50	2.46
G. S.	1.30	2.37	.74	.77
J. B.	.40	.74	.19	.47
N. P.	.69	1.14	.28	.56

This difference measure is used rather than the absolute maximum decision-time, so that our comparisons will be independent of the base value of decision-time. The base value ('immediate' decision) is something which depends upon the reaction-time of the individual (as determined by previous attitude and the instructions) in a way with which the present theory is not concerned.

It will be remembered that the 'Accuracy' condition is assumed to increase the restraining forces over that of the 'Usual' condition;

[1] The answer to the question of when the correspondence between data and theory is close or is not close must be made on a relatively arbitrary basis. No *a priori* rules can usually be laid down for such a judgment. The particular situation will determine whether inconsistencies in the data and deviations from the predictions can be considered unimportant or are sufficient to warrant the conclusion that the theory is not corroborated. In the following presentation of results the author has stated his own judgments in this respect.

that the 'Speed' condition is assumed to lower the restraining forces below that of the 'Usual' condition; and that the 'Constant' condition is assumed to increase the restraining forces in one direction, thus increasing the average restraining force over that of the 'Speed' condition. (Since we do not know to what extent the restraining force will be increased in the 'Constant-Error' condition, we can only compare it to the 'Speed' condition in this case.)

Table 3 shows that in every case the results are in line with the theory. The difference between the highest and lowest decision-time points is always greater for the 'Accuracy' condition than for the 'Usual' condition; it is always less for the 'Speed' condition than for the 'Usual' condition; and it is always greater for the 'Constant-Error' condition than for the 'Speed' condition.

Derivation 3 seems to be corroborated without exception.

Derivation 4

The precision of the relative frequency ogive increases as the restraining force increases. (Judgment becomes most accurate as the restraining force increases.)

This derivation may be alternatively stated: the sigma of the relative frequency ogive decreases as the restraining force increases—since the sigma is inversely related to the precision. Table 4 presents

TABLE 4

MEANS AND STANDARD DEVIATIONS IN STIMULUS UNITS OF THE RELATIVE FREQUENCY OGIVE

Subject	Usual		Accuracy		Speed		Constant	
	Mean	Sigma	Mean	Sigma	Mean	Sigma	Mean	Sigma
H. S.	+2.58	1.62	+2.10	1.21	+2.17	1.55	+ .83	1.41
B. W.	+1.68	1.22	+2.47	1.19	+2.82	1.28	+1.70	1.25
G. S.	+3.56	1.63	+3.04	1.52	+2.23	1.48	+ .43	1.38
J. B.	−2.00	2.73	−2.44	2.58	−1.41	2.38	+1.10	2.15
N. P.	+1.08	1.62	+ .84	1.65	+ .72	1.56	+1.73	1.30

the sigmas of the relative frequency ogives for each *S* under each condition.

A glance through this table under the columns headed 'sigma' shows another factor to be operating. The general trend is for the sigma to decrease in the successive conditions. The more time the *S*s spent making judgments in the experimental situation, the more accurate they became. This is not predicted from the theory, but is easily understandable. The mere process of becoming accustomed to the situation, being more and more at ease with the procedure and

being able to adjust one's perception to the demands of the situation, may very easily lead to the observed result of increasing accuracy of judgment. This disturbing factor prevents an adequate test of the derivation.

There are, however, deviations from this decreasing trend, and these deviations seem to be in line with the predictions. Thus, subject H. S. and subject B. W. both show sigmas which increase from the 'Accuracy' to the 'Speed' condition, in line with the theoretical predictions. The sigmas should, according to the theory, decrease from the 'Usual' to the 'Accuracy' conditions and again from the 'Speed' to the 'Constant' condition. In these cases, therefore, the effect of adapting to the situation is working in the same direction as the effect of the change in restraining force. With these two factors supplementing each other, it is not surprising that four of the five *S*s show sigmas which decrease from the 'Usual' to the 'Accuracy' condition (subject N. P. shows a negligible increase), and all five show sigmas which decrease from the 'Speed' to the 'Constant' condition.

It is rather equivocal as to whether or not Derivation 4 is supported by the data. On the whole, considering the rather strong trend for the sigmas to decrease, and the exceptions in line with the theory, one would be inclined to accept Derivation 4 as tentatively corroborated.

Derivation 5

If the restraining force is of different magnitude in different directions, the 50 percent point of relative frequency is shifted towards the side of the greater restraining force.

Subjects H. S., B. W., and G. S. had been told in the 'Constant' condition that they had previously been making an error in the direction of saying 'shorter' too many times. For these *S*s, then, the restraining force against "saying 'shorter'" is greater than against "saying 'longer.'" According to the theory the 50 percent point of the relative frequency ogive should shift in the direction of 'shorter.'

Table 4 gives the means of the relative frequency ogives calculated with the use of the Müller-Urban weights. These values give the position of the 50 percent point on the stimulus continuum, the positive direction being that in which the variable is longer than the standard, and the negative direction being that in which the variable is shorter than the standard. Thus, a value of + 2.58 means that the 50 percent point of the relative frequency ogive occurs when the variable is 2.58 sixteenths of an inch longer than the standard.

The results are in all cases according to the theoretical prediction. In the case of subject H. S., the mean of the ogive in the first three

conditions was above + 2.1. In the 'Constant' condition it shifts to + .83, an appreciable shift toward the 'shorter' direction. Subject B. W., in the 'Usual' condition, has a mean of + 1.68. For the 'Accuracy' and 'Speed' conditions which precede the 'Constant' condition, however, the mean of the ogive is well over + 2.0. In the 'Constant' condition it shifts back toward 'shorter' to + 1.70. For subject G. S., also, the 50 percent point of the relative frequency ogive in the 'Constant' condition is much farther over toward 'shorter' than in any of the three previous conditions.

Subjects J. B. and N. P. were told in the 'Constant' condition that they had been saying 'longer' too many times. From these two *S*s, then, we would expect 50 percent point of relative frequency to shift in the direction of 'longer' in the 'Constant' condition. Subject J. B., who initially had the 50 percent point in the negative direction, shifts past the point of objective equality, and the 50 percent point for the 'Constant' condition is at + 1.10. Subject N. P. shows a mean under the 'Constant' condition of + 1.73, while previously its greatest value had been + 1.08.

We may regard Derivation 5 as being strongly supported by all five *S*s.

We may then conclude that the gross predictions from the theory have been rather well supported by the obtained data.

Summary

An experiment, specifically oriented toward empirically testing some derivations from a theory of decision, is described. The derivations, together with the data bearing upon them, have been presented. The data, on the whole, are sufficiently in line with the theoretical expectations for them to be accepted as corroborative evidence.

Evidence that an ogival relationship exists between confidence ratings and stimulus differences is presented and discussed. This relationship is important in the experimental test of the more precise quantitative predictions from the theory.

(Manuscript received November 2, 1942)

References

1. Cartwright, D. Decision-time in relation to differentiation of the phenomenal field. *Psychol. Rev.*, 1941, **48**, 425–442.
2. Cartwright, D. The relation of decision-time to the categories of response. *Amer. J. Psychol.*, 1941, **54**, 174–196.
3. Cartwright, D., & Festinger, L. A quantitative theory of decision. (To be published in *Psychol. Rev.*)
4. Festinger, L. Studies in decision: II. An empirical test of a quantitative theory of decision. (To be published in this journal.)
5. Johnson, D. M. Confidence and speed in the two-category judgment. *Arch. Psychol.*, 1939, **241**, 1–52.

TORSION IN PERSONS WITH NO KNOWN EYE DEFECT

BY THOMAS G. HERMANS

University of Washington

The problem of torsion, the rotation of the eyeballs about their sagittal or antero-posterior axes, arose in a discussion of the horopter as stimulated by Carr's "Introduction to Space Perception." The horopter is "the locus of all points in the binocular field of vision, the images of which fall upon identical points of the two retinas" (**6**). The description of the horopter is quite dependent upon the nature of torsional movements that might occur. Carr (**6**) comes to no conclusion regarding torsion, but states "some authors conclude that the eyes exhibit opposed torsion in convergent movements. It is said that the upper parts of the eyes rotate toward each other in changing the fixation from a distant to a near position, and rotate away from each other in divergent movements." On making various tests with my own eyes, it appeared that Carr's authorities are probably mistaken as to the direction of torsion occurring with changes in convergence. On referring to contemporary authorities ample reason is found for Carr's inability to make any conclusive statement regarding the nature of torsion either in direction or amount. Briefly, physiologists say that it does not occur in normal vision, and opthalmologists regard it as abnormal.

The history of the problem of torsion dates back to the controversy over the laws of Listing and Donders. These laws have provoked a good many experiments on, statements of opinion about, and techniques devised to determine the nature of torsional movements. The experiments have been sketchy and have yielded data of no statistical significance; the opinions expressed are confused and influenced by individual peculiarities in vision and variation in introspective skill; the differences in techniques have given rise to conclusions which may be artifacts of the technique used. The resulting current opinion from supposedly authoritative sources is, as one might expect, confused.

Current Opinion

In current physiology texts, discussions of torsion occur in conjunction with the description of the functioning of the external eye muscles. The authors describe the individual reaction of each muscle and how it works in conjunction with and in opposition to others, but point out that the combined action is such as to prevent torsion. Starling (**11**) says in speaking of the oblique muscles, "The function of these two small muscles appears to be to prevent the eyes from rotating

about their visual axes, and in particular to prevent the rotation inwards which is associated with the contraction of the superior and inferior recti. For this purpose the superior rectus is associated with the inferior oblique, and vice versa. In this action the oblique muscles appear to be very efficient; for if the eye is first fatigued by looking at a brilliant line of light, *e.g.*, a long straight electric lamp filament, and is then directed upwards or downwards at a white surface, the after-image thus produced is always found to keep its vertical direction." Best and Taylor (3) say, "It will be seen . . . that the obliques and the superior and inferior recti when contracting individually produce a rotary or wheel-like movement. When acting in pairs the rotary actions being in opposite directions antagonize one another so that normally no such movement occurs." Howell (8) mentions the torsional effect of individual muscles but does not mention torsion in any combined action of the muscles. MacLeod (9) also implies that the recti and oblique counteract each other in preventing torsion, and summarizes, "Under normal conditions the coordination of the eye muscles is *extraordinarily exact*." [1]

Turning to other sources than physiology texts we find the situation not quite so hazy on the problem of torsion. Duke-Elder (7) says, "It is evident . . . that starting from the primary position the cardinal movements (vertically and horizontally) are not associated with torsion, but any oblique motion is accompanied by a *certain amount* of torsion. This forms the basis of the physiological definition of the primary position of the eyes, as that position from which vertical and horizontal movements are not associated with torsion." By using the after-image technique (we will pass for the moment the errors of observation resulting from its use) he reports that with oblique positions up and to the right and down and to the left the torsion is to the right; that with oblique positions up-left and down-right the torsion is to the left. "Further, no matter how any position has been obtained, whether by the eyes directly towards it from the primary position, or by moving them first along the horizontal direction and then along the vertical, or by a quite irregular route, the after-image will be found to occupy exactly the same position in every case. It can be shown, moreover, that the movements of the two eyes are *precisely similar*, for if either is closed the after-image as seen by the other occupies exactly the same position."

Bielschowsky (4) says, "A torsional deviation takes place only in those extreme diagonal positions of the eyes which do not occur in ordinary vision because owing to compensatory movements of the head all extreme movements of the eyes are avoided." Further on, however, he says in describing various combinations of muscles to effect certain movements, "In a fourth combination the vertical motor muscles effect a rotary movement in opposite directions, namely a conclination (intorsion) or disclination (extorsion) of the vertical meridians. The innervation required for that rotation is *brought about only* through corresponding fusional stimuli, for example, that presented by two identical objects which are fused in the stereoscope or the haploscope and are rotated in opposite directions. The eyes follow this displacement within certain limits, so that the images will remain on corresponding retinal areas, and binocular single vision will thus be maintained."

Ames and Gliddon (1) give an account of torsion which is typical of the clinician's approach to such a problem. They measure cyclophoria (torsion) in clinical cases "suffering more or less acutely from their eyes." They speak of "cyclodeclination condition," "cyclotorsional trouble," "abnormal cyclotorsional condition," and "suffering from cyclodiplopia" as if torsion were an abnormality of vision, yet they make no reference to what normal persons, or non-clinical cases might show in the way of torsion. But further, they state, "Usually the declination of the vertical meridians is greater than that of the horizontal. The amounts of both vertical and horizontal declinations usually vary with convergence. Furthermore, in many cases it has been found that the difference between the declinations of vertical and horizontal meridians varies with convergence."

Carow (5) reports measurement of torsion in 107 employees of a Berlin police station. This is the only source found in which there is any attempt to establish norms for torsion in healthy persons. Most of his subjects had eye defects of one kind or another, but at least they were not recruited from clinics. Only 14 yielded what he regarded as good data—"*Gute Angaben.*" What his criterion of 'good' might have been is not clear. However, he cannot be accused of labeling as good only those showing torsions consistent with some bias, for his averages for the whole group show torsions to approximately the same extent and direction that the 14 showed. The averages he gives, which constitute his attempt at statistical analysis of the data, are quite consistent with the findings herein to be reported.

---

[1] Italics here and in quotations which follow are mine.

There have been published only two reports of experimental attempts in American psychological laboratories to settle the problem of torsion. Barnes (2) observed directly torsional movements in two subjects by use of a telescope containing cross-hairs which could be superimposed upon distinctive markings of the iris. The amount of torsional change for two positions of the eye could be read from a vernier scale on the outside of the telescope. Due to ambiguities in reporting her findings, the account is practically worthless. For this reason, Loring (10) repeated the experiment using three subjects. Her analysis of the data collected makes her conclusions rather nebulous. She simply counts the number of right torsions as compared with left torsions, and finds "a predominance of torsion of the left eye to the left, and of the right eye to the right, from the point of view of the experimenter," and further "a tendency for the torsion to increase as the eye passes from the primary position to successively more oblique positions, including horizontal, vertical and 45 degree meridians."

Methods of Testing Torsion

There are three essentially different methods of testing torsion. It is my belief that some of the conflicting statements about the phenomenon arise as an artifact of the method used. One method frequently recommended uses the after-image. This method is complicated by the fact that the surface upon which the after-image is projected must be always at right angles to the visual axis; if a continuous surface is used, it would have to be spherical and equidistant from the eye at all points. Furthermore, only monocular after-images should be used to study torsion. If torsional movements were in the same direction, clockwise or counter-clockwise, and the same amount for both eyes, the binocular after-image would show such torsion; but if the torsional movements were bilaterally symmetrical or in any way dissimilar, the perceived position of the after-image would tend to remain a function of the identical areas stimulated during the retinal fatigue, and might be perceived as remaining in the same position regardless of the angle of projection. The statement previously quoted from Starling resulted, presumably, from the use of the binocular after-image, although he used the singular of 'eye.' Furthermore, the use of the binocular after-image leads to the probably erroneous (at least for oblique positions) statement illustrated in the quotation from Duke-Elder "that the movements of the two eyes are precisely similar, for if either is closed the after-image as seen by the other occupies exactly the same position." Tscherning (12) makes the same error in observing the nature of torsion. That these observations are probably erroneous is indicated by observations, such as those of Loring, of each eye in oblique positions. One eye may undergo a different amount or direction of torsion, or both, than the other when the two assume a position of oblique fixation. Yet the binocular after-image would not show validly the torsion in either one or both, due to the fact that a binocularly fixated point or figure, seen as single during fixation, never produces a double after-image.

A second method is that of direct observation of the iris by means of a telescope; this method seems desirable because of its objectivity; the experimenter is not dependent upon getting subjects who are good introspectionists. However, there are certain limitations to its use. The telescope used by Loring is still in the local laboratory. On experimenting with it I found difficulty in discriminating with any degree of certainty changes of less than a degree. Loring took the average of 10 readings for each position, and probably arrived at a fair estimate. The telescope is also difficult to use with subjects who have dark eyes, because on a dark background the cross-hairs are not readily discernible. Loring reports that pupillary changes even with constant illumination made it difficult at times to identify the distinctive marking previously sighted. A further difficulty with this method of observing the single eye is that of controlling convergence when attempting to measure torsional changes as they occur with change in the angular position of the eye. As will be shown later, there are torsional changes with changes in convergence. Barnes and Loring note that convergence should be controlled but fail to indicate how this was accomplished. The advantage of this method in addition to its objectivity is that torsion may be observed with lateral angular displacement of the point of fixation.

A third method of measuring torsion is that of presenting to the two eyes separate figures which are binocularly fused when their angular rotations on the visual axes are in proper relationship. If the ordinary stereoscope is used, however, convergence cannot be varied and it is difficult to vary the visual angle to positions other than points in the medial plane. A device for stereoscopic presentation of figures with which convergence can be varied is described by Ames and Gliddon (1). It utilizes the mirror stereoscope principle with the mirrors in tubes that rotate about axes at the interpupillary distance. With this device the convergent positions of the eyes are always for points in the medial plane, but it could be rigged to test torsion with lateral fixations. It should here be noted that with the stereoscopic method relative torsion only is being measured and not the torsion of the eyes individually. It is legitimate to assume that torsions measured thus with medial fixations result from an equal amount of torsion in both eyes—that the binocular movements are bilaterally symmetrical. We may assume also that with lateral or oblique fixation points relative torsions would not be attributable to equal torsions of the two eyes. If this is true, we would want to know the individual torsions of the two eyes and not relative torsion; so it is pointless to toy with stereoscopic techniques enabling lateral fixations.

A feature not to be neglected with any of these stereoscopic devices is the nature of the figures or targets to be binocularly fused. The

type of targets commonly used is what has come to be known as 'Volkmann's disks.' These are circles set at right angles to the visual axes with the centers as points of fixation; a radius is drawn on each circle, but on opposite sides of the centers, so that when one or both disks are properly rotated the radii become an apparent diameter of the circle. The angular deviation from actual linearity that is necessary to make a diameter, were there no torsion, indicates the amount of relative torsion. The task for the subject or observer is one of reporting when the radii appear to make one continuous straight line. If each circle has a diameter drawn on it, the task then becomes one of discriminating, while fixation is centered on the circle, doubleness at the ends of the diameter. Of course, with diameters, fixation may be at any point along the fused line, but the doubleness to be discriminated is *always peripheral* to the point of fixation. If the two-point discrimination threshold is proportional to the peripheral distance from the fovea, the angular size of the target in the visual field would make no difference in measuring relative torsion. If, however, the two-point discrimination is an absolute amount independent of foveal distance, the larger the target the more precise the determination of relative torsion. Rather than two-point discrimination threshold this should be called threshold of discrimination of dissimilarity of retinal points.

The point at issue in evaluating what Ames and Gliddon (1) report concerning cyclofusional amplitude is this problem of threshold of discrimination of dissimilarity of retinal points. "The cyclofusional amplitudes show the capacity of the eyes to maintain fusion of print or other objects, which are rotated about the point of fixation in a plane perpendicular to the visual axes." This idea of cyclofusional amplitude implies the truth of what Bielschowsky (4) says as previously quoted, *i.e.*, that torsion occurs as a pursuit motion to maintain a single image. I doubt that such pursuit torsions occur; at least there is no proof of it, and the "Cyclofusional amplitude of Mr. Ames' eyes" in the table given might equally well be labeled "Mr. Ames' discrimination ability." They give figures showing cyclofusional amplitudes which increase with increase in convergence. This fact seems at first more consistent with the notion of pursuit torsions than with apparent cyclofusional amplitude being dependent upon discrimination ability. However, if one considers the fact that there is a decrease in perceived size with increase in convergence, it is not surprising that these cyclofusional amplitudes increase with convergence; it is plausible that discrimination ability decreases with decrease in apparent size. Very strong evidence discrediting Bielschowsky's pursuit torsion notion is given in the following demonstration which anyone may try on himself. Fuse two vertical straight lines stereoscopically. These being parallel, the resulting fusion will be perceived as vertical and straight. Now, tip the lines gradually toward each other and then away from each other in a continuous movement. The resulting stereoscopic depth perception is that of a straight line being tilted toward the eyes and away from the eyes in a continuous movement. If Bielschowsky's pursuit torsions occurred, this perceived movement would be discontinuous within certain middle limits of rotation on either side of vertical, with the depth effect suddenly appearing in the perception beyond those critical limits of the pursuit torsions. It is this nondiscriminated doubling of the retinal images that is the essence of stereoscopic depth perception. If the doubling is too great the person perceives two lines coming together at an angle. It would seem that this limit of nondiscriminated doubling is what Ames and Gliddon are reporting as cyclofusional amplitudes.

Returning to the nature of targets, Ames and Gliddon (1) report the use of radii targets in determining cyclodeclinations (torsion) in the horizontal as compared to the vertical. They find, "Usually the declination of the vertical meridians is greater than that of the horizontal."

This unexplained peculiarity of the eyes' torsional movements might be related to the vertical-horizontal illusion. Or could it not be that this generalization might arise from errors in sampling? This is suggested by the fact that the only values given are those in Table 6, "Cyclodeclinations of Mr. Ames' Eyes" in which the reverse of the generalization is indicated. They question the use of radii on the targets for the reason that they would be testing the "cyclorelationship, not between corresponding parts of the two retina, but between parts that lie on opposite sides of the fovea"; they believe that there might exist an asymmetry in the two retinas, and the method used should "make it possible to compare the cycloposition of corresponding parts of the two retinas." They describe a target they use consisting of three small concentric circles marking the center as a point of fixation and radiating beyond the circles in vertical and horizontal meridians three dots on one target, and on the other, in identical positions, small holes through which light shines from a source behind the target. The circles fuse and the dots and light points superimpose when the targets are rotated the proper amount about their centers. These targets are essentially diameters requiring peripheral discrimination while fixation is upon the fused center, which is difficult for a naïve subject. Within groups of college students I usually have found 20 to 40 percent reporting they had never observed before that the image of an object is double when the point of binocular fixation is at a different distance than the object. One would expect little validity in data obtained on torsion with use of diameter targets that require peripheral discrimination of double images.

From the foregoing discussion of current opinion and evaluation of methods used in determining torsion, it becomes apparent that many questions about torsion remain without definite answers. 1. Is torsion a normal or an abnormal phenomenon? 2. If it occurs in normal vision, what direction of torsion may be expected? 3. What is the average amount that might be expected? 4. What is the normal variability that might be expected from one person to the next? 5. Does it vary with variation in elevation of vision above and below horizontal? 6. Does it vary with changes in convergence? 7. Are the torsional movements bilaterally symmetrical for the two eyes? 8. What are the torsional changes with oblique and lateral fixation points in the visual field? 9. Is it a source of distress in vision?

The following report of experimental procedure and the results obtained is an attempt to answer all but the last three of these questions.

Apparatus and Techniques of Measurement

From a careful consideration of the various methods of measuring torsion, it seemed advisable to use the stereoscopic technique with targets of the Volkmann disks type, particularly in view of the fact I had had previously made for research on another problem a well constructed and accurately calibrated telestereoscope.[2] Fig. 1 shows the essential parts of the telestereoscope. The four-legged base and the frames for mounting the mirrors were cast in bronze and tooled afterward. The use of cast bronze minimized the necessity of seasoning after casting to allow for warping; thus tooling could proceed immediately. The mirrors are placed perpendicular to the base and the outer ones can be driven synchronously through a range of 40 degrees by means of the crank and gears. The gear ratio is such that one complete turn of the crank, *8*, turns each mirror one degree. Due to the double mirror relationship, a one-degree turn of the outer mirrors will necessitate a change of two degrees in each eye for continued binocular fixation of a point. The

[2] A note of appreciation is due Mr. George E. Price, shop technician for the local laboratory, for the excellent work done in constructing the apparatus used in the present study, and the assistance rendered in making the drawings of the graphs and figures in the present article.

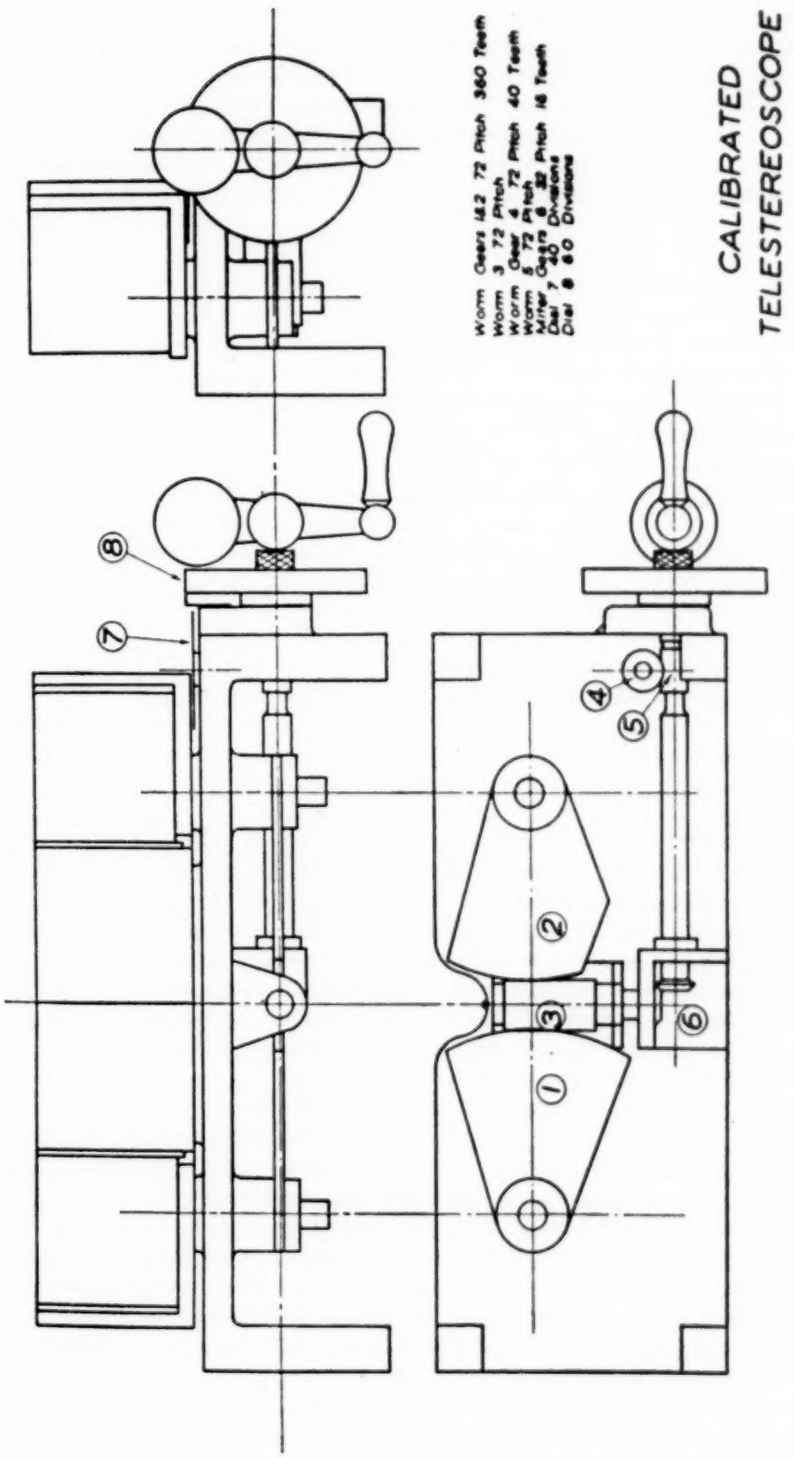
CALIBRATED
TELESTEREOSCOPE
Worm Gears 42 72 Pitch 360 Teeth
Worm 3 72 Pitch
Worm Gear 4 72 Pitch 40 Teeth
Worm 5 72 Pitch
Miter Gears 6 32 Pitch 16 Teeth
Dial 7 40 Divisions
Dial 8 80 Divisions
1
2
3
4
5
6
7
8

FIG. 1

crank wheel, *8*, is calibrated for minutes of angular change of each mirror while the degrees are indicated on the horizontal dial, *7*, above. The worm drive, *3*, is tapered 1/100 in. to allow for tightening the play of the gears when wear has occurred. The mirrors are made of thin plate glass.

The telestereoscope was mounted at one end of a box 4 ft. long, 14 in. wide, and 11 in. deep, inside the other end of which the targets were placed with the mechanism for turning them on the outside of the box. A partition running the length of the box separated the right and left visual fields. The interior was painted a dull black which effectively eliminated everything within it from the subject's visual field except the white targets; these were illuminated by ten-watt globes shielded from vision at the near end of each alley. The mechanism driving the targets is shown in Fig. 2. The axes of rotation of the targets at *1* and *2* are the centers of circles whose segments

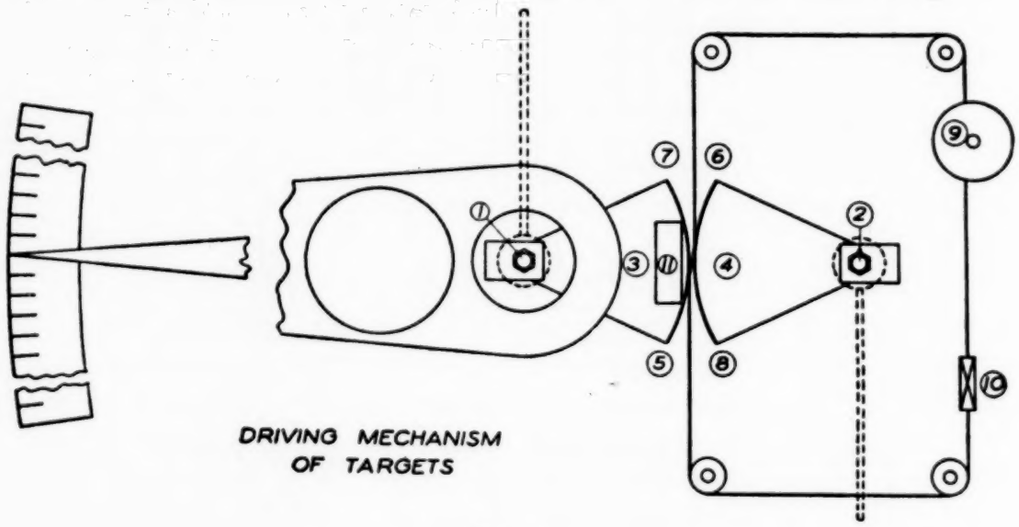

FIG. 2

3 and *4* are made of sheet metal. A cable connecting *5* and *6* pulls against a cable whose ends connected at *7* and *8* is driven by the subject turning a knob, *9*, over the axle of which the cable is wound. This cable arrangement makes possible synchronous but opposed rotary motion of the targets, when the subject turns the knob on the outside of the box with his left hand. The pointer rotating on axis 1 is the length of the radius of a circle whose circumference is 360 cm. Thus, the scale value of one cm. indicates one degree of rotation of the target and readings can be taken to the nearest mm. or tenth of a degree. The pointer is made of sheet metal with edges folded to increase rigidity and is counter balanced with lead, *11*, soldered to the circle segment *3*. Circular sections were cut out of the pointer arm to reduce its weight.

The targets for stereoscopic fusion were circles 1 in. in diameter. Projecting 4½ in. beyond the circles were bars 1/8 in. wide; with one circle the bar projected downward and with the other upward. The targets are shown in Fig. 2 in broken lines. In effect, these targets combined binocularly as radii of a circle 10 in. in diameter, and appeared when properly adjusted as a straight line passing through the 1 in. circle. The axes of rotation of the targets were placed at a distance apart that would meet the requirements for the assumed minimum interpupillary distance of 58 mm. In other words the targets were placed so that the person with pupils 58 mm. apart would be fixating the targets with parallel axes of vision when the telestereoscope was set for zero degrees of convergence. It was calculated that with an assumed maximum interpupillary distance of 68 mm., placement of the targets at this distance apart would require convergence of less than ½ degree, when the telestereoscope was set at zero.

The box was hinged to a frame at a position approximating the center of vertical rotation of the subject's eyeballs when his forehead was pressed against a stereoscope frame in a stationary position in front of the mirrors of the telestereoscope. This placed the hinges (the center of vertical rotation of the box), the subject's eyes, and the targets at all times in the same plane. The remainder of the frame was rigged with stops at 5 degree intervals from 45 degrees below to 45 degrees above horizontal.

On using this apparatus in testing vision, one could expect that at some combination of the two variables, elevation and convergence, average values of torsion would approximate zero. On

preliminary testing of 23 subjects, averages were obtained which showed a constant deviation from zero. This indicated a possible error in the apparatus, either in the setting of the targets in parallel lines with the pointer at zero on the scale, or in the position of the mirrors. By use of a surveyor's transit the error was found in the mirrors. Viewing the target through the transit directly the lines were found to be vertical when the pointer indicated zero on the scale; but when viewed with the transit through the mirrors the targets had to be rotated between 7 and 8 tenths of a degree to make the reflected lines vertical. No closer estimation of the error was attempted and since this would be a constant error for all positions of the mirrors all values obtained were corrected (arbitrarily) .7 degree.

Experimental Procedure

Preliminary testing of 23 volunteer students from my classes indicated the necessity of obtaining subjects who had had recent eye examinations and had been passed as 'normal'; that is, with so-called 20–20 vision and no detected defect. The best source of such material seemed to be the naval R.O.T.C., a highly selected group of men. Through the coöperation of the officers [3] of this unit, I was fortunate in obtaining 104 subjects with the required visual standards and willingness to coöperate. Each subject was told that he was to adjust the targets so that they made a continuous straight line; that there was no correct response, but that I merely wanted to know at what position of the targets the line looked straight to him; that, however, for each eye position two judgments would be made, the second to serve as a check on the first, and that too great a difference in pairs of judgments would indicate that he was not attending to what he was supposed to be doing. On the first disagreement in the two judgments of more than a degree he was stopped and shown the position of the targets at the positions of the two judgments and asked if he could not discriminate a difference, which he could. He again was assured that each of his judgments was right but that he should be more consistent. He was not stopped on further inconsistencies. There was no distinguishing feature on the knob the subject turned in moving the targets that might assist him in duplicating his previous judgment.

The following routine in the subject's manipulation of the target was insisted upon: each judgment started by throwing the targets out of line on both sides of where the line made by them might appear to be straight and ending the judgment by narrowing down the amplitude of movement in an active process, rather than one of turning the targets to where the line appeared almost straight and then staring at it until it looked straight. There were good reasons for insistence upon this method of moving the targets. The most obvious reason is that movement is a factor determining attention and thus the judgments should be more consistent when the method was followed. Another reason is that fixation of a bent line, such as the targets would make when placed at an angle, causes the line after a time to appear less bent, and the after effect is to cause a line actually straight to appear bent in the opposite direction.

An hour was allowed for the running of each subject on the test; the length of time required was dependent upon the subject's degree of abulia and the number of successive steps in covergence he could make without the target appearing double. Each was started with the box in a horizontal position and with the telestereoscope mirrors set for parallel axes of vision. At this position some found it impossible to fuse the targets, in which case the mirrors were changed to one degree of convergence; only a few found fusion impossible at this position. Beyond 2 degrees of convergence some who had been able to fuse the targets with parallel axes of vision found it impossible to do so. Each subject continued with successive steps in convergence until he reported fusion impossible. (The degree of convergence at which fusion ceased was quite consistent at all elevations for a given subject, and only a few gave any evidence of what might be regarded as learning the trick of converging.) Following the horizontal position of the box, testing proceeded with elevations of 10, 20, 30, and 40 degrees above and below horizontal. To cancel possible fatigue effects, half the subjects proceeded with elevations above horizontal first, and half proceeded downward. Each subject was told to maintain the same head position as when looking horizontally and make his eyes do the turning up or down. Vertical shifts in the head position were necessary for extreme positions of elevation; this was accomplished by changing to a stool of different height; also, the subjects were encouraged to rest their eyes by looking out the window

[3] I am very grateful to Capt. Eric L. Barr and Lt. Com. Carlisle Thompson for their coöperation in recruiting volunteers to serve as subjects in this experiment, and grateful to the volunteers for their alert and interested performance as subjects.

and by rolling their eyes whenever it seemed desirable. For these reasons head clamps or a biting board for maintenance of constant head position would be a nuisance and seemed to be a needless precaution, particularly in view of the fact that I sat at the subjects' right, within reach of the telestereoscope, and could see readily and check any tilting of the head to compensate for extreme positions of elevation.

Analysis of the Data

A distribution of the averages of the subjects' two judgments for each of the 99 combinations of convergence and elevation was made, and the mean and standard deviation from the mean calculated for each distribution. The values derived from these calculations are given in Table 1; the horizontal variable represents degrees of convergence as indicated by the calibrations on the telestereoscope, and the vertical variable, the degrees of elevation of vision above and below horizontal. The first item in each cell of the table indicates the number of subjects able to make judgments at this combination of the two variables. The second item is the mean of the distribution in degrees, corrected for the aforementioned error in the mirrors' placement. The direction of rotation of the targets is a tipping of vertical diameters of the fused circles of the targets away from each other at their tops. Correspondingly, the rotary direction of the eyes is that which would result from unopposed contraction of the inferior oblique muscles. For simplicity in speaking of torsion we shall assume that these figures refer to torsion of the eyes; for obviously, the eyes have undergone torsion of these amounts and the targets have had to be rotated in the same amount and direction to meet the subject's requirements for judgments of straight lines. The direction of the torsions indicated (with two exceptions) is an ex-torsion or disclination of the upper ends of the vertical meridians of the eyes from the medial plane. The two exceptions in the table as indicated by the minus sign occur at -40 degrees in elevation and at 0 and 7 degrees of convergence. The third item in each cell is the standard deviation from the distribution's mean. The fourth item is the theoretical mean; how this was obtained will be explained later. All values were calculated to four decimal points; the values in the tables are given to the nearest .01 degree.

Attention should be called here to the difference in the degree of torsion found in this study and what might be obtained in other studies, as dependent upon the technique of measurement employed. The targets used in the present study were driven synchronously in opposite directions of rotation; the values derived here refer to the amount of rotation of each target necessary to make the binocularly fused target appear to be crossed by a continuous straight line; the assumption is that each eye makes torsional movements of this amount. If only one target were rotated we could expect values

TABLE I

GIVING: (1) NUMBER OF SUBJECTS, (2) OBTAINED MEAN VALUE OF TORSION, (3) STANDARD DEVIATION, (4) THEORETICAL MEAN, FROM 40° ABOVE TO 40° BELOW HORIZONTAL AND FROM 0 TO 10 DEGREES OF CONVERGENCE.

	0°	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°
+40°	77	95	95	87	72	55	40	30	23	14	10
	.16	**.43**	**.55**	**.71**	**.81**	**1.02**	**1.19**	**1.51**	**1.74**	**2.24**	**1.95**
	.74	.78	.79	.85	.75	.86	.79	1.02	1.13	1.77	1.09
	.20	*.33*	*.48*	*.64*	*.81*	*.99*	*1.18*	*1.38*	*1.60*	*1.83*	*2.06*
+30°	84	103	101	95	87	67	49	37	31	25	15
	.05	**.19**	**.30**	**.39**	**.54**	**.74**	**.78**	**1.14**	**1.54**	**1.82**	**2.00**
	.57	.55	.65	.57	.66	.66	.79	.81	.98	1.04	.80
	.16	*.28*	*.41*	*.56*	*.71*	*.88*	*1.06*	*1.25*	*1.45*	*1.66*	*1.88*
+20°	85	102	103	97	93	78	52	37	31	27	22
	.04	**.18**	**.30**	**.41**	**.51**	**.70**	**.88**	**.96**	**1.25**	**1.50**	**1.63**
	.41	.40	.49	.49	.52	.56	.70	.66	.72	.64	.72
	.13	*.23*	*.35*	*.48*	*.62*	*.77*	*.93*	*1.11*	*1.29*	*1.49*	*1.70*
+10°	87	103	104	100	89	72	50	40	32	27	21
	.10	**.16**	**.26**	**.38**	**.46**	**.60**	**.79**	**.92**	**1.09**	**1.31**	**1.63**
	.41	.43	.42	.49	.52	.54	.51	.51	.55	.65	.58
	.10	*.19*	*.29*	*.41*	*.53*	*.67*	*.81*	*.97*	*1.15*	*1.33*	*1.52*
0	90	103	104	101	89	71	51	42	33	28	15
	.07	**.17**	**.23**	**.35**	**.43**	**.49**	**.72**	**.91**	**1.08**	**1.08**	**1.28**
	.30	.36	.38	.48	.52	.56	.59	.57	.46	.62	.71
	.07	*.14*	*.23*	*.33*	*.44*	*.56*	*.70*	*.84*	*1.00*	*1.17*	*1.34*
−10°	84	102	104	100	94	79	58	46	38	30	19
	.13	**.19**	**.29**	**.41**	**.51**	**.62**	**.65**	**.85**	**1.04**	**1.18**	**1.23**
	.43	.43	.46	.53	.58	.61	.61	.61	.66	.70	.66
	.04	*.10*	*.18*	*.26*	*.36*	*.46*	*.58*	*.71*	*.85*	*1.01*	*1.17*
−20°	81	102	104	101	94	80	59	43	37	27	15
	.10	**.14**	**.22**	**.30**	**.44**	**.44**	**.57**	**.70**	**.85**	**.88**	**.83**
	.44	.44	.46	.56	.57	.55	.64	.50	.52	.59	.54
	.02	*.06*	*.12*	*.19*	*.27*	*.36*	*.47*	*.58*	*.71*	*.85*	*1.00*
−30°	79	102	102	99	91	78	59	43	26	19	14
	.03	**.06**	**.15**	**.23**	**.29**	**.37**	**.34**	**.41**	**.59**	**.42**	**.54**
	.47	.45	.43	.54	.56	.57	.79	.87	.81	.58	.44
	−.01	*.03*	*.07*	*.12*	*.19*	*.27*	*.36*	*.46*	*.57*	*.69*	*.83*
−40°	69	99	104	101	92	82	64	42	32	20	11
	−.07	**.01**	**.09**	**.15**	**.21**	**.26**	**.17**	**−.04**	**.16**	**.12**	**.13**
	.51	.55	.51	.58	.66	.79	.79	.95	.69	.77	.59
	−.03	*−.01*	*.02*	*.06*	*.11*	*.17*	*.25*	*.33*	*.43*	*.54*	*.66*

twice this amount. To illustrate, let us compare the values for torsion given by Carow (**5**) with those obtained in the present study. He used a target for the left eye consisting of stationary graph paper with vertical and horizontal lines, and for the right eye a movable disk with a diameter drawn on it. In fusing the two, the diameter was moved to coincide with or appear parallel to the vertical lines of the graph paper. The values to be expected from this technique of measurement should be approximately twice those of the present

results. He measured torsion when points of fixation were at remote, 20 cm., and 15 cm. distances, which correspond approximately to the 0, 9, and 10 degrees convergence of the present study; his variations in elevation were +20, 0, and −20 degrees. Table II shows the

TABLE II

COMPARISON OF CAROW'S TORSION VALUES WITH THOSE OF THE PRESENT STUDY

	0°	9°	10°
+20	1.45	3.70	4.92
	.13	1.49	1.70
0	1.13	2.71	3.58
	.07	1.17	1.34
−20	.77	2.16	2.73
	.02	.85	1.00

comparison of Carow's averages with the theoretical values of the present study; the latter are the lower items of the cells. The direction of torsion in Carow's study is the same as in this study, but the amounts are consistently greater than twice the value of mine. This suggests the possibility of a constant error in his apparatus analogous to the one discovered and corrected for in mine, particularly in view of the fact that none of his average values approximates zero.

Fig. 3 shows the three-dimensional surface obtained when the obtained mean values of torsion in degrees are plotted against degrees of elevation and degrees of convergence. The striking appearance of regularity of this surface suggests the desirability of obtaining an equation descriptive of the relationship between the three variables. The quadratic surface:

$$Z = ax^2 + by^2 + cxy + dx + ey + f$$

was fitted to the data by the method of least squares.[4] This surface was selected on the assumption that no true points of inflection are to be found in the data. In fitting this surface, the relative reliability of the different means was allowed for by weighting each squared error, between the obtained and expected values of the dependent variable, inversely as the sigma of the mean on which it was based, *i.e.*, directly as the square root of the number of cases. The solution was carried through by means of the Doolittle method. The equation becomes:

$$Z = +.00573x^2 + .00001y^2 + .00148xy + .07029x + .0028y + .06876$$

in which Z is degrees of torsion, x degrees of convergence, and y elevation of vision in degrees above or below (plus or minus) horizontal.

[4] I am grateful to Dr. William Z. Birnbaum, Professor of Mathematics, and Mr. Douglas H. Lawrence, University Fellow in Psychology, for advice in carrying out this process.

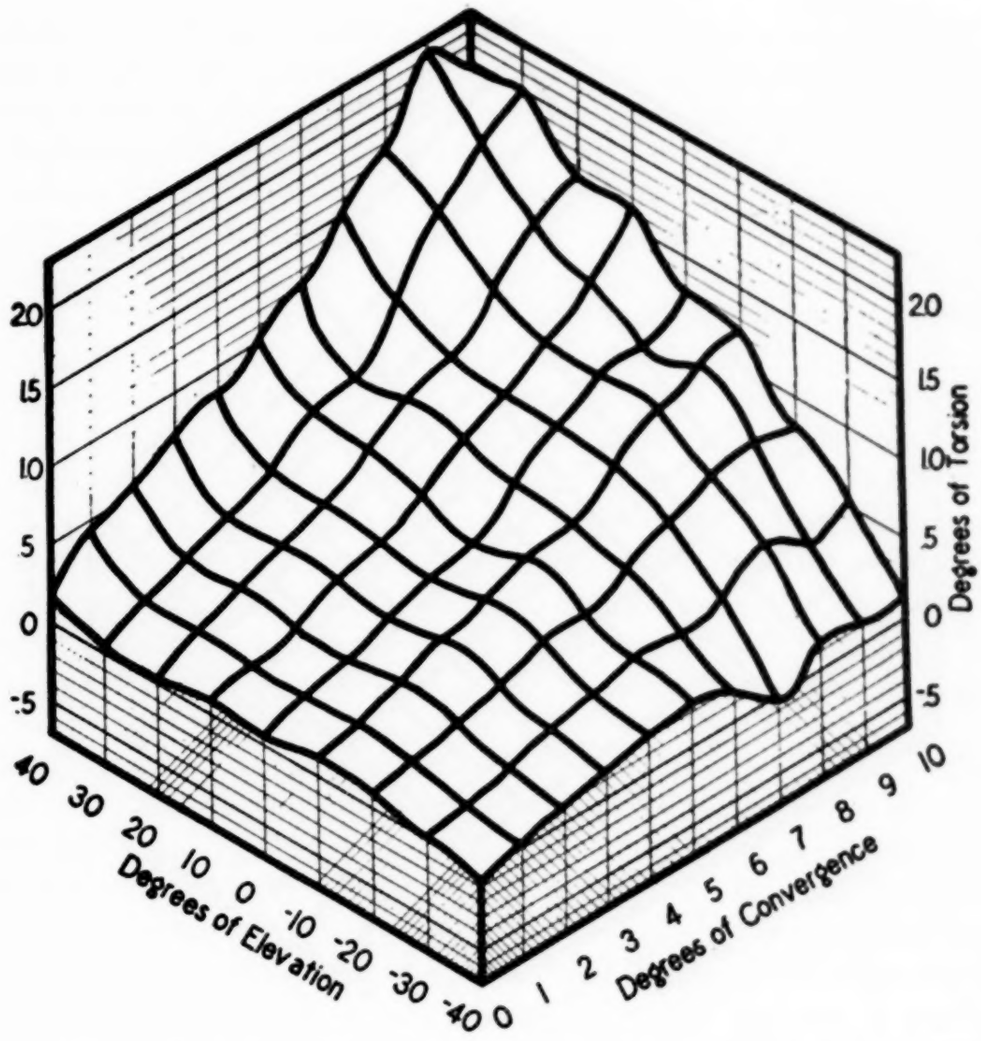

FIG. 3. Obtained mean values of torsion plotted against the two variables, convergence and elevation of vision.

Using this equation, theoretical values of torsion were calculated for the 99 points of the graph. The fourth item in each cell of Table I is this theoretical value. Fig. 4 shows this smoothed surface. The sigma of the differences between the obtained and theoretical means is .13 degrees.

DISCUSSION OF RESULTS

On inspecting the data of the present experiment several questions arise. One may wonder at the wide variation in the number of subjects at the different points of variation of convergence and elevation. With the telestereoscope set for parallel axes of vision one may expect some subjects, due to variation in interpupillary distance, to find it difficult to fuse the targets stereoscopically because of inability to diverge their visual axes. However, due to the placement of the targets to meet requirements for an interpupillary dis-

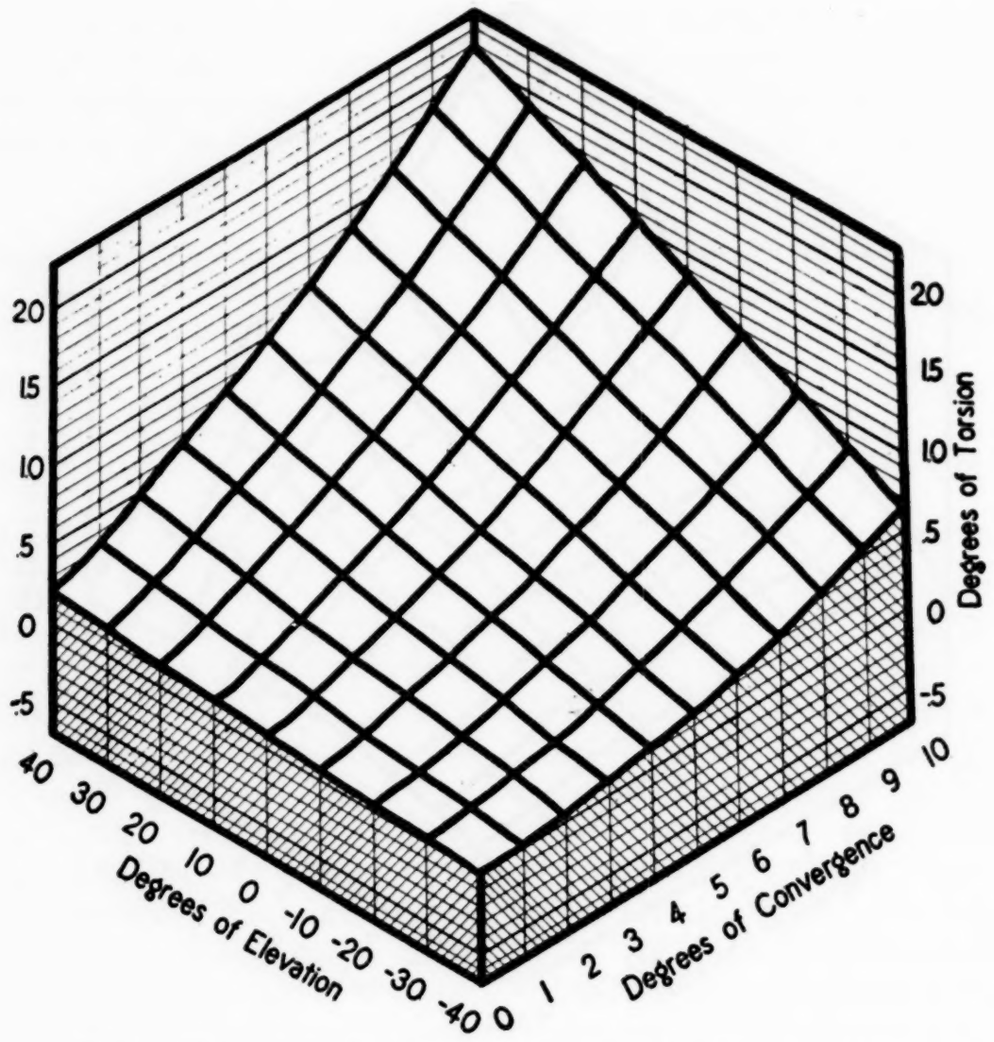

FIG. 4. Theoretical values of torsion plotted against the two variables, convergence and elevation of vision.

tance of 58 mm., practically no one found divergence necessary at this setting of the apparatus. The average interpupillary distance of all the subjects was 63.76 mm., standard deviation 2.99; the average of the 41 persons finding it impossible to fuse the targets at some position of zero convergence was 63.36 mm., S.D. 2.51, while the average was 63.96, S.D. 3.24 for the others. Apparently a small interpupillary distance is not a major factor in this inability to fuse the targets when the instrument is set at zero, although the difference in averages points in that direction. The difficulty in fusing the targets seems to be due to the inability to dissociate accommodation and convergence; this is no doubt true especially at positions of extreme convergence, for most subjects reported increasing indistinctness of outline of the target with increase in convergence, and, incidentally, a decrease in apparent size, and an apparent approach of the targets upon increase in convergence. Some referred to my

manipulation of the crank of the telestereoscope as moving the targets toward them, not knowing that convergence was increasing.

One wonders, also, if the subjects who are able to fuse the targets with extreme convergence are representative of the whole group in the amount of torsion shown. If they were not, one would expect better than chance correlation between the amount of torsion shown at positions of little or no convergence and the number of convergent positions at which the subject was able to fuse the targets. Such correlations were calculated at the following positions chosen randomly, el. + 40 degrees—conv. 1 degree, el. + 20 degrees—conv. 0 degrees, el. + 10 degrees—conv. 2 degrees, and el. − 40 degrees—conv. 2 degrees; the correlations, respectively were + .17 ± .06, + .02 ± .07, + .05 ± .06, and − .13 ± .06. Apparently there is no significant relation between the amount of torsion and the number of convergent positions. From this fact we can assume that the values of torsion obtained with the few on extreme convergence are representative of what might be expected if all the subjects had been measured with that degree of convergence.

Another point to be questioned is the reliability of the individual subjects. This may be indicated by the average disagreement within the subject's pairs of judgments. Distributions of these differences for the first 50 subjects were made in an effort to find justification for rejecting the data obtained with certain subjects whose torsions did not agree with my biases. For example, the marked irregularity in the obtained data at el. − 40 degrees—conv. 7 degrees is produced by the abnormality in torsion of one subject; if it were possible to show that he was erratic and unreliable in his judgments it might be legitimate to throw out his data. However, he was no more inconsistent in his judgments than others who could converge that far; in fact it became apparent that the larger the degree of convergence the greater the disagreement in the pairs of judgments that might be expected. This is probably due to the apparent decrease in size with increase in convergence and also to the blurring of the target caused by changes in accommodation consistent with convergence, but inappropriate for the objective distance of the target. It became all the more apparent that to exclude the data of any subject would be violating one aim of the present study, *i.e.*, to determine the variability in torsion that one might expect in persons with supposedly normal vision. Returning now to the nature of the distributions of the subjects' agreement with themselves on the pairs of judgments, all the distributions were skewed, with the mode at about .2 degree, the median less than .5 degree, and the mean varying with the number of convergent positions at which the subject could fuse the targets.

Any further or more precise statistical analysis on this point seemed pointless.

Also, any analysis of the curves for torsional values of individual subjects is pointless, and only leads to that kind of thinking on the problem, so richly illustrated in its history, that this study is attempting to throw into the discard. The individual curves showed remarkable regularity and similarity to the curves given by the averages. Furthermore, some subjects showed negative torsions or conclinations of the vertical meridians relative to the medial plane, but this is resident in the size of the standard deviations, *i.e.*, the limits on both sides of the average value within which the middle 68 percent of the values of torsion fell.

Two important questions arise as to the meaning of the data of the present study in relation to certain almost classical problems in vision. The first pertains to the nature of the horopter, and the second pertains to the concept of primary position and Listing's Law. In the definition of the horopter cited at the beginning of this paper we are playing safe, but the definition commonly encountered in psychology texts describes it as an imaginary surface within which all points are seen as single with binocular vision. If the theoretical surface of Fig. 4 is valid as a description of torsion, the horopter is a surface at only a few points on this surface, and at all other points it is only a line in the medial plane slanting away from or toward the observer dependent upon the degree and direction of torsion at that point; the horopter is a surface only when torsion is zero.

If one recalls the already mentioned source of error in the apparatus, one may well question the arbitrary correction of the obtained values .7 degree. Data were recorded and all calculations and graphs made before the correction was applied, and since the error of the mirrors could be determined by use of the transit only as being between .7 degree and .8 degree, one may ask why .8 degree or some other value was not selected. Several sources have contended that zero torsion occurs with parallel axes of vision between 10 and 30 degrees below horizontal, the most commonly used position of the eyes in reading and walking about. This seems entirely plausible, and since a .7 degree correction fits the data better in this theory, it was used. If a correction were made for the placement of the targets at 58 mm. interpupillary distance rather than the average 63 mm. for the group of subjects, zero torsion would be closer to horizontal than the .7 degree correction places it. The question of where zero torsion might be is not be decisively answered.

Warren (**13**) defines 'primary position' as the position which the eyes assume when the head and body are erect and the eyes fixate an infinitely distant point in the medial and horizontal planes. This

definition is meaningful and useful, but is not the one commonly used in the literature on Listing's Law and torsion. The quotation from Duke-Elder (7) previously cited gives the concept of primary position commonly used; "the physiological definition of the primary position of the eyes," is "that position from which vertical and horizontal movements are not associated with torsion." On the basis of the present findings it would seem there is no primary position in the sense of this definition. Warren gives Listing's Law as "a principle of eye movement formulated by Listing which states that if the eye moves from the primary position to any other, the torsional rotation of the eyeball in this new position is the same as it would be if the eye had turned about a fixed axis, lying at right angles to the initial and final directions of the line of regard." In either sense of 'primary position' Listing's Law has validity for a very limited range of eye movements, namely, only for movements between two points of zero torsion. These lie below horizontal (probably) and within a narrow range of convergence.

Conclusions

1. Torsion is a normal phenomenon of vision.

2. The direction of torsion to be expected is a disclination at the top of the vertical meridians of the eyes relative to the medial plane, except at extreme positions of elevation below horizontal and with a small amount of convergence; in this position the torsion is reversed.

3. The average amount of torsion to be expected for any combination of convergence and elevation is expressed in the equation,

$$Z = +.00573x^2 + .00001y^2 + .00148xy + .07029x + .0028y + .06876$$

in which Z is degrees of torsion, x degrees of convergence, and y degrees of elevation above (plus) or below (minus) horizontal; the standard error of estimate of this equation is .13 degrees.

4. The amount of variability in torsion that one might expect in individual cases is expressed best by the various standard deviations from the average values as indicated in Table I.

5. The direction of progressive change in torsion with increase in elevation (convergence constant) is a disclination at the top of the vertical meridians of the eyes relative to the medial plane; this is greater, the greater the degree of convergence.

6. The direction of progressive change in torsion with increase in convergence (elevation constant) is also such a disclination; this is greater, the higher the elevation.

The present study has shown that Listing's Law has very limited validity, and establishes norms in support of Donders' Law, which is

"a principle of visual fixation, according to which every position of the lines of regard in relation to the head corresponds to a definite, invariable angle of torsion of the eyes, regardless of the path by which the position has been reached" (**13**). Furthermore, it is shown that the horopter can be only a line except within a very limited range of positions wherein torsion is zero.

(Manuscript received October 23, 1942)

References

1. Ames, A., & Gliddon, G. H. Ocular measurements. Tr. Sec. Ophth. A.M.A., 1928. Pp. 102–175.
2. Barnes, B. Eye-movements. *Amer. J. Psychol.*, 1905, **16**, 199–207.
3. Best & Taylor. *The physiological basis of medical practice.* (2nd ed.) Baltimore: William and Wilkins Co., 1940, p. 1651.
4. Bielschowsky, A. Disturbance of the vertical motor muscles of the eyes. *Arch. Ophthal.*, 1938, **20**, 175–200.
5. Carow, R. Untersuchungen über die Näherungsrollung und die Hebungs-Senkungsrollung der Augen. *v. Graefes Arch. Ophthal.*, 1939, **140**, 86–115.
6. Carr, H. A. *An introduction to space perception.* New York: Longmans, Green and Co., 1935, p. 190.
7. Duke-Elder, W. S. *Text-book of ophthalmology.* St. Louis: C. V. Mosby Co., 1939, p. 594. Vol. 1.
8. Howell, W. H. *A textbook of physiology.* (14th ed.) W. B. Saunders Co., 1940. Pp. 358–359.
9. MacLeod's *Physiology in modern medicine.* Edited by Bard and others. (8th ed.) St. Louis: C. V. Mosby Co., 1938, p. 230.
10. Loring, M. W. An investigation of the law of eye-movements. *Psychol. Rev.*, 1915, **22**, 354–370.
11. Starling's *Principles of human physiology.* Edited by Evans and Hartridge. (7th ed.) Philadelphia: Lea and Febiger, 1936, p. 340.
12. Tscherning, M. *Physiologic optics.* Philadelphia: The Keystone, 1904, p. 292.
13. Warren, H. C. *Dictionary of psychology.* Edited by Howard C. Warren. Houghton-Mifflin Co., 1934.

THE REVERSAL OF DISCRIMINATION IN A SIMPLE RUNNING HABIT

BY R. N. BERRY, W. S. VERPLANCK AND C. H. GRAHAM

Brown University

Our experiment is concerned with the 'reversal' of a discrimination in an experimental situation similar to one used by Verplanck (**6**). The term 'reversal' refers only to the conditions of reinforcement and non-reinforcement, and is used for convenience in description. When a reversal takes place in the present experiment, the originally reinforced stimulus is made the unreinforced stimulus and the unreinforced stimulus is made the reinforced stimulus. In this descriptive sense, the first set of conditions under which the rat learns the problem constitutes the original learning. The first change in the conditions of reinforcement is called the first reversal, and the second change is called the second reversal. Actually, the second reversal reinstates the conditions of original learning.

Evidence from experiments by Hunter (**3**) and McCulloch and Pratt (**4**) indicates that an original discrimination interferes with the establishment of the reversed discrimination. On the basis of results obtained on the lever-pressing apparatus, Skinner (**5**) concludes that induction effects are present between the reinforced and non-reinforced stimuli in the first reversal but not in the second. Some results of Fritz (**1**) show no carry-over of effect from an original discrimination to its reversal.

Our learning situation presents two measures, latent period and running time. The measures are most easily understood in terms of the experimental situation which gives rise to them. A full description is given by Verplanck (**6**). The latent period is defined as that period between the time at which the door in front of the starting box (of the 'runway apparatus') is opened and the time at which the rat, with the exception of tail, crosses a mark three inches in front of the starting box. The second measure, running time, is the time from the moment the rat crosses the mark to the moment when the door in front of the food box is dropped behind the animal.

The latent period may be thought of as a measure of the 'running out' response to the stimulus of opening the door. Running time may be considered as a measure of the 'running-down-the-alley' response. On the basis of this analysis we may think of the two measures as representing the respective strengths of two *S-R* members within a chain of correlations [*cf.* Verplanck (**6**)].

METHOD

The apparatus is identical with the one used by Verplanck. One side of the runway is painted black; the opposite side is white; the sides are gray. By this arrangement the animal runs over a runway which presents two discriminative stimuli: 'white' or 'black.'

Original discrimination for the rats utilized in this experiment was established by Verplanck. Verplanck employed 4 groups of rats in his experiment. Rats from two of these groups were employed in the present research. Both of these groups (Verplanck's Groups I and IV) were originally trained to make the black-white discrimination. Each rat in Group I had run 6 cycles daily, each cycle consisting of (a) one reinforced trial on the white alley and white food box and (b) three trials on the black alley and black food box. Each rat in Group IV had received the same training except that he was run for only 4 cycles daily. All rats, when they came to the present experiment, had received a total of 96 trials on the actual discrimination problem. This number does not include the preliminary training and acquisition trials before discrimination was introduced.

On the day that the reversal of discrimination was introduced, the rats were first run for 3 retention cycles. Following these cycles and on the same day, the rats ran for 3 cycles on the new discrimination problem. On the 3 succeeding days, 6 cycles per day were run. On the fifth day, data were collected for three cycles. All together, over a period of five days, the first reversal involved a total of 96 trials.

The first 3 cycles of the second reversal were instituted in the second half of the fifth day's experimental session. During the next 4 days, 96 trials were completed by each rat on this reversal. Six cycles (or 24 trials) were run each day by each rat, an exception being the last day, on which only 3 cycles were completed.

A complete trial consisted of the following sequence. The rat was placed in the starting box. As soon as the experimenter was ready, the door in front of the starting box was raised, so that the rat might run down the alley into the food box. If the rat entered the food box, the food-box door was dropped behind him. The animal was permitted to remain in the box and eat for one minute and 25 seconds. At the end of this interval, the food box was picked up and carried around to the other end of the apparatus. The rat was placed tail foremost in the starting box. Thirty-five seconds after the food box had been picked up, the door in front of the starting box was again opened, and another trial was begun. In the case of reinforced trials, the runway was turned over before the rat was taken back to the starting box, and food was placed in the food box. In the case of unreinforced trials, the runway was rattled and the can from which the food pellets were taken was moved about. This procedure was introduced so that the rat might receive no discriminative stimuli from these sources.

If, during the trials, the latent time was more than two minutes, the starting box door was closed and the next trial was started 30 seconds later [*cf.* Graham and Gagné (2) and Verplanck (6)]. In case the running time exceeded two minutes the rat was replaced in the starting box, and the starting box door raised 30 seconds later. The two-minute period was used because previous work with this experimental apparatus indicated that if the rat did not run or enter the food box within this period, he was not likely to do so at all. Times were taken by means of two stop watches.

The animals used in this experiment were male members of the inbred Wistar strain which constitutes the colony of the Psychological Laboratory. All rats had already been trained on the original discrimination by Verplanck. Each of the two of Verplanck's groups which were used was further divided. A subgroup, made up of rats from Groups I and IV, was run on the first retention trial 3 days after Verplanck's last trial of original discrimination. The remaining rats in Groups I and IV were tested 21 days after the original discrimination. Table I indicates the number in each of the 4 groups thus created.

TABLE I

NUMBER OF RATS IN EACH GROUP DURING THE EXPERIMENT

	Group I	Group IV
3 day retention	5	4
21 day retention	3	3*

* One rat in this group had a 29 day retention period, instead of the 21 day interval.

All of the rats were maintained in good health throughout the period from the original learning to the retention trials by relaxing the feeding rhythm whenever it was considered necessary. However, all rats were returned to the twenty-minute-per-day feeding rhythm two days before the new problem was begun. During this regime the rat was fed a total of 20 minutes a day, part of which was spent in the apparatus.

Results

Although the rats were divided into the 4 groups of Table I, the results, as finally presented, are computed for all the data rather than for the particular groups. So far as can be seen, no significant group differences appear in the results of the first and second reversals. It appears that differences due to the original discrimination and the conditions of retention are too small to have any effect upon later training. For this reason it seems justifiable to average the data on all the groups, at least for purposes of arriving at a first approximate description of reversal. The records of original conditioning for the rats utilized in this experiment were taken from Verplanck's data.

All measurements have been converted into logarithms in accordance with the procedure used by Graham and Gagné (2) and by Verplanck. Graham and Gagné have discussed the reason for using logarithms in this situation. The measure of central tendency employed is the median. (The logarithm of the median is, of course, the same as the median of the logarithms.) It is impossible to obtain a significant mean because the rat might fail to leave the starting box or to enter the food box. In these circumstances, times (running time and latent period) are 'infinite,' and by their indeterminate nature cannot be used in the computation of a mean. However, some sample means on trials showing no 'infinite' times are in very close agreement with the medians calculated for the same trials. Verplanck found this same agreement between the median and the mean. No measure of variability is reported. Deviations are of the same order as those obtained by Verplanck.

The results [1] of the experiment are presented in Figs. 1 and 2. Fig. 1A presents the medians of log latent period for all the unreinforced trials. Retention trials and the points of reversal are indicated in this figure and also in those which follow. Each of the values in this figure represents the median of all the unreinforced trials in any given cycle. Fig. 1B presents the medians of log latent period for the reinforced trials. Data on running time are given in Figs. 2A and 2B.

The log latent period curve for unreinforced trials (Fig. 1A) on the original discrimination differs somewhat from the one obtained by Verplanck. This is due to the lumping of the two groups in the

[1] Numerical data are contained in a thesis filed in the John Hay Library of Brown University (R. N. Berry. The reversal of discrimination in a simple locomotor habit, 1941.)

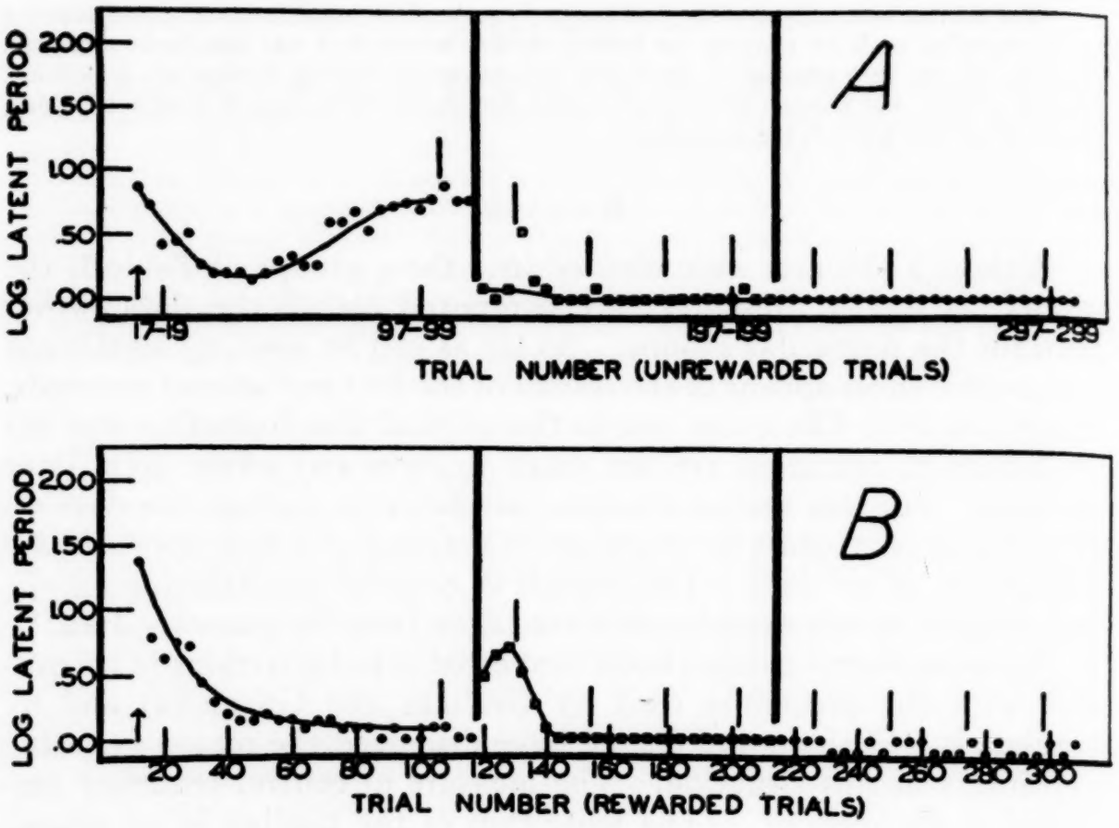

FIG. 1

A. The course of response (measured by latent period) to the unreinforced stimulus.
B. The course of response to the reinforced stimulus.
In both graphs the heavy vertical lines separate original discrimination from the first reversal and the first reversal from the second. The arrow indicates the trial on which discriminative stimuli were first introduced. The three points through which the smooth line is not drawn represent retention trials. The short vertical lines above the curves separate daily runs.

present experiment, to the smaller number of animals, and, most importantly, to the fact that we have not attempted to plot the daily extinction curves. Rather, the over-all trend is indicated, and this is true not only for the first discrimination but also for the succeeding reversals. Verplanck's data show significant trends which are concealed when Groups I and IV are combined because these two groups were not run the same number of trials per day. The curve for original discrimination shows a decreasing latency followed by a rising latency until the latter approaches a final level. Values for the retention trials after original learning are indicated in the last three points, through which the curve does not pass.

The curve of the first reversal, separated from original discrimination by a vertical heavy line in Fig. 1A, shows a low initial value, a rise to a maximum, and a decline to a final level. The low initial value might be expected since this response had been the reinforced one on the retention trials and during original discrimination. The

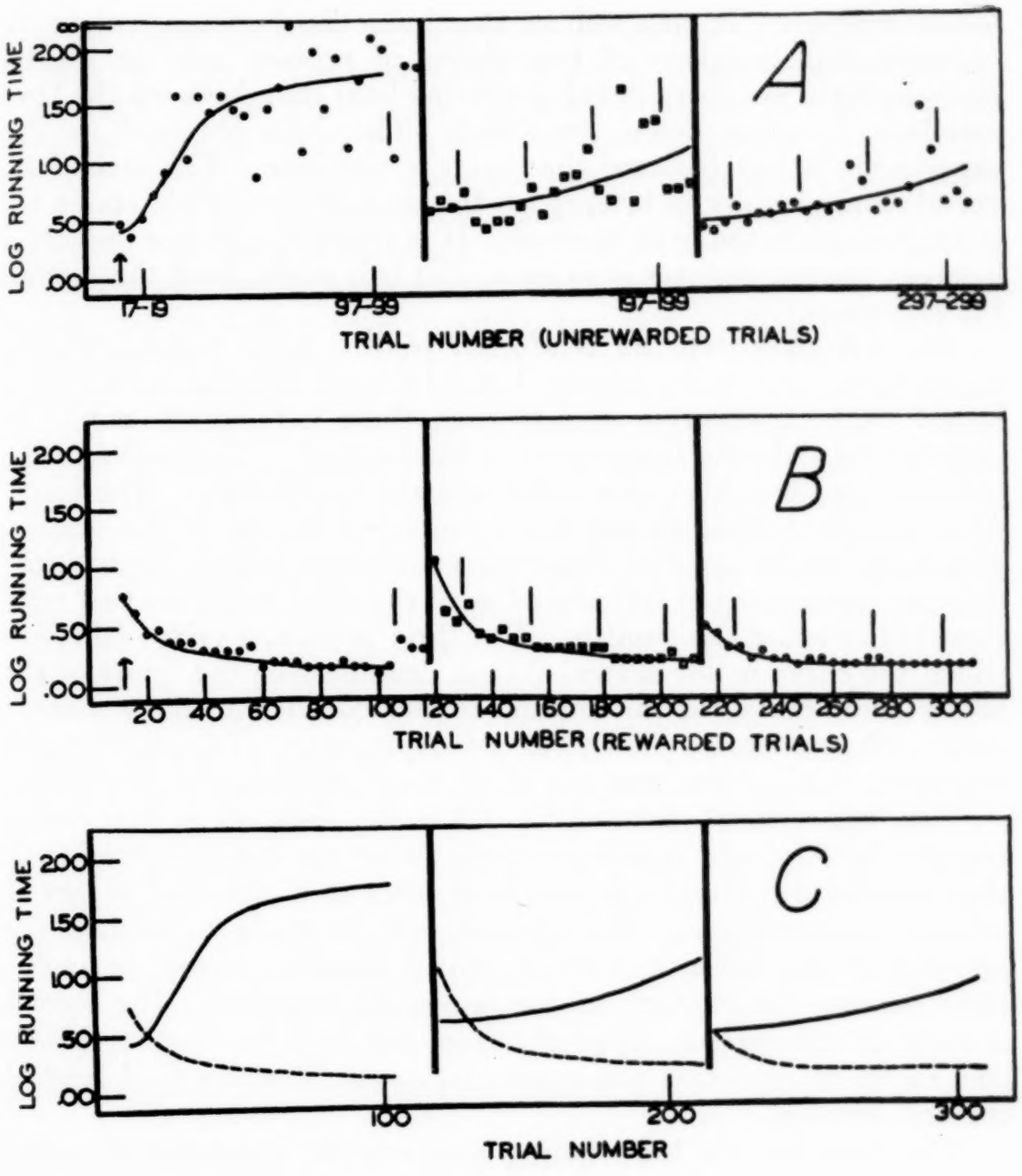

FIG. 2

A. The course of response (measured by running time) to the unreinforced stimulus. Symbols as in Fig. 1.

B. The course of response to the reinforced stimulus. Symbols as in Fig. 1.

C. Curves of Figs. 2A and 2B placed on the same graph for purposes of comparison. Solid line: responses to unreinforced stimuli. Dashed line: responses to reinforced stimuli.

first value on the second day of the reversal is noticeably above the values for the first day. It is possible that the non-reinforcement of the day before is having more effect than the reinforcement which preceded in retention and in original discrimination. The sharp drop which follows this single high value might be attributed to an increasing amount of induction between the reinforced and unrein-

forced reflexes.[2] It may well be that induction between temporally corresponding members of two chains of reflexes may be greater between early members (leaving starting box) than between the later members (running down the alley). The data obtained in this experiment would indicate that such is the case. The second reversal shows no rise in latency on the unreinforced trials above the last latencies in the first reversal. It is possible that the induction between the two reflexes is so great that it is maintained throughout the experiment.

Fig. 1B gives data for reinforced trials. Since Verplanck's log latent period curves for Groups I and IV were identical on the reinforced trials, a selection of rats from these two groups might be expected to yield the same curve as he obtained. This proves to be the case, and it is his curve which is fitted to the data. The first 8 trials run by Verplanck are not represented in any of the graphs. This is due to the fact that these trials constitute original acquisition; discrimination was not introduced until the ninth trial, the first reinforced trial being the twelfth. The three retention cycles (through which the curve is not drawn) do not deviate from the limits of the curve. The curve for the original reinforced trials exhibits a rapid drop and decelerates slowly to an asymptote. This final value represents the limiting response under these experimental conditions.

The first reversal (data of Fig. 1B to the right of the first heavy vertical line) for the reinforced trials shows an initial increase and then a rapid drop nearly to the level reached in the final stages of original discrimination. No attempt will be made to explain the presence of this initial rise which clearly appears to indicate an obvious decrease in strength of the reinforced response. The second reversal in this graph, separated from the first by another heavy line, shows no deviation throughout its course from the final level of the first reversal.

The data for the log running time on the unreinforced trials, shown in Fig. 2A, exhibit great variability. Verplanck has shown that this variability is only apparent, and that the data are fitted by smooth extinction curves drawn through the records for single days. However, we have not attempted to analyze original discrimination in this way. Indeed, such an analysis would be impossible, since the

[2] In the light of the consistently short latent periods on both reinforced and unreinforced trials after the first reversal, it was thought possible that the animal might be responding to the opening of the door and not to the alley at all. If so, he might give a response great enough to cause his falling from the starting box in the absence of the alley. To test this possibility, the alley arrangement was reconstructed so that the alley might be removed from the apparatus. When the latent times had reached a low level in the course of the first and second reversals, the alley was removed at the beginning of a few unreinforced trials. At no time did an animal fall or nearly fall out of the starting box when the alley was absent.

data are compounded of scores for two groups varying in the number of cycles per day. The curve, then, is drawn to represent a tendency rather than a rigid description. It is important to observe that the data for retention trials (the last three points before the curve for the first reversal) show no apparent deviation from the final values of the curve for original discrimination. This indicates that the change in experimenters does not influence the level of discrimination. The curve for the first reversal indicates a slow, but unstable, increase in the running time for the unreinforced trials. The second reversal shows the same general characteristics as the first. Perhaps the increase from the first trials to the last in the second reversal is slightly smaller than the corresponding increase during the first.

The original learning curve for log running time on reinforced trials in Fig. 2B fits the curve obtained by Verplanck, and it is his curve which is used. It shows a slow deceleration through the first four cycles and beyond this point attains a final level. The three retention cycles show slightly higher values than the last ones of original learning, but the differences do not appear to indicate lack of retention to any significant degree.

The initial section of the first reversal in Fig. 2B exhibits a higher initial value than original discrimination. This is probably due to the fact that the animal is responding to the reinforced stimulus as if it were still unreinforced. The curve probably does not reach the low terminal value that is reached in original discrimination, but if we disregard the value for the first cycle, the shape of the curve seems to be essentially the same as that for original learning.

The second reversal has an initial value lower than that of the initial value of either original discrimination or first reversal. The terminal values of this curve are as low as those in original discrimination and perhaps lower than those in the first reversal.

Discussion

The general trends of the curves in Figs. 2A and 2B can be better understood if they are plotted on the same graph, as in Fig. 2C. The graph facilitates visual comparison. Little is to be gained by making a similar graph for the log latent period data (Figs. 1A and 1B). The latter values reach a constant level for both reinforced and unreinforced responses soon after the first reversal, and this level is maintained throughout the course of the experiment.

In considering Fig. 2C we may assume that the height of the curve for reinforced trials, at any ordinate value, represents some inverse measure of the reflexes: running down the alley to (a) the reinforced stimulus and (b) the unreinforced stimulus. The differ-

ence between corresponding ordinate values, then, is some measure of the difference in strength between the reinforced and unreinforced responses at a particular stage of conditioning. In these terms, we may say that the strengths of the reinforced and unreinforced responses show a smaller difference at the end of the first reversal than they do at the end of original discrimination. A similar diminished difference exists between the two reflex strengths at the end of the second reversal. Further, it may be observed that the difference in strengths of the two reflexes increases more slowly during the first and second reversals than it does during original discrimination. The first reversed discrimination, then, develops neither so rapidly nor so strongly as the original discrimination, and this is true of the second reversal, even though this reversal reinstates the conditions of the original discrimination. We have seen this effect in greatly augmented form in log latent period data. Soon after the first reversal the differences in strength measured by log latent period drop nearly to zero.

The term induction implies that two related reflexes mutually interact so that each is modified positively or negatively in strength by the other, depending on the characteristics of the influencing reflex. In the present experiment inductive effects seem to be present (a) between the reinforced and unreinforced reflexes occurring early in the chain (latent period measures) and (b) between the reinforced and unreinforced reflexes occurring late in the chain (running time measures). Induction is particularly strong in the early member, where the latent period measure is shown to reach a relatively constant value, uninfluenced by reinforcement or non-reinforcement, early in the first reversal. Since the latency finally attained is at or near the final level of latency to the reinforced stimulus during original discrimination, this must mean that almost the full weight of induction is in the direction: $R_{pos.}$ to $R_{neg.}$

The same effect is shown in the reflexes measured by running time, although to a lesser extent. Here the differential response to the unreinforced stimulus is clearly maintained, but in diminished degree, as contrasted with its value in the original discrimination. On these grounds it may be true that induction more seriously influences an early member of a chain than a later one. If so, the explanation probably lies in the relative strengths assumed by the member correlations at their respective temporal (or spatial) intervals from reinforcement. Thus, in the present experiment, the extinctive effect of non-reinforcement has too little influence to counteract the excitatory induced effect from preceding reinforcements of the positive responses as measured by latent period. This is not true in the running time measures of the later members. Here, the extinctive effect is close

in time and space to the final unreinforced response; hence, the response is less influenced, in absolute degree, by the induction effect from the reinforced response.

To the extent that different measures allow of comparison, our results seem to be in agreement with those obtained by Hunter (3), McCulloch and Pratt (4), and Skinner (5).

Summary

The present experiment has been concerned with the courses of (a) original discrimination, (b) its reversal (first reversal), and (c) its reinstatement (second reversal). The method employed involves the acquisition of a discrimination by a white rat on the runway apparatus of Graham and Gagné (2). During original discrimination, the animal's response is reinforced on the white alley and unreinforced on the black alley. Ninety-six trials on the original discrimination are followed by 12 retention trials, 3 or 21 days later. Conditions of reinforcement and non-reinforcement are then reversed (first reversal). The second reversal starts 96 trials after the first trial of the first reversal and lasts for 96 trials. The second reversal reinstates the conditions of original discrimination.

The course of discrimination has been studied by means of measures of two members of the chain of reflexes leading to entrance into the food box. These two measures are *latent period* (early member) and *running time* (late member).

1. Latent period for the reinforced response during original discrimination falls from a high value and eventually reaches a final steady level. Latent period for the unreinforced responses shows an over-all rise to a relatively unstable final level.

2. With the introduction of the first reversal, the latent period exhibits a brief rise from the terminal level of original discrimination but soon declines to a low final value. The unreinforced response measured by latent time shows a significantly high value on the first full day of reversal, but thereafter it soon reaches the low level characteristic of the reinforced response.

3. During the second reversal, latent periods for both responses are maintained at a steady low level comparable to the final level reached by the latency of the reinforced response during original discrimination.

4. Running time (the measure of the later member of the chain) on reinforced trials, throughout all the discriminations, starts at a high value and thereafter drops to a steady final value. On unreinforced trials, running time increases during each discrimination. This increase is greatest during original discrimination.

5. The decreased difference between running times on reinforced and unreinforced trials, characteristic of the first and second reversals, is interpreted in terms of increased induction processes. This interpretation is also applied to the data on latent period.

(Manuscript received December 2, 1942)

References

1. Fritz, M. F. Long time training of white rats on antagonistic visual habits. *J. comp. Psychol.*, 1930, **11**, 171–184.
2. Graham, C. H., & Gagné, R. M. The acquisition, extinction, and spontaneous recovery of a conditioned operant response. *J. exp. Psychol.*, 1940, **26**, 251–280.
3. Hunter, W. S. Habit interference in the white rat and in human subjects. *J. comp. Psychol.*, 1922, **2**, 29–59.
4. McCulloch, T. L., & Pratt, J. G. A study of the presolution period in weight discrimination by white rats. *J. comp. Psychol.*, 1934, **18**, 271–290.
5. Skinner, B. F. *The behavior of organisms.* New York: D. Appleton Co., 1938, pp. ix + 457.
6. Verplanck, W. S. The development of discrimination in a simple locomotor habit. *J. exp. Psychol.*, 1942, **31**, 441–464.

THE ROLE OF AUTISM IN A VISUAL FIGURE-GROUND RELATIONSHIP*

BY ROY SCHAFER AND GARDNER MURPHY

College of the City of New York

The formal dynamics of the figure-ground relationship as developed by the Gestalt school have wide relevance in relation to perception, judgment, thought, and the attitudinal structuring of social situations. But we are urgently in need of more information regarding the figure-ground dynamics of perception.

Simple, standard procedures in the study of visual perception make a good starting point. The objective situation in terms of lines and shading can be definitely known and personal variables controlled. Shifts of organization and emphasis can be clearly defined.

The factor in visual perception with which we are concerned is the influence of autism; autism is defined here as the organization of cognitive processes in the direction of *need satisfaction.* Four studies may be mentioned as a background for the present experiment. The two pioneer studies by R. N. Sanford (3, 4) showed the influence of food-need upon interpretation of ambiguous figures. The third, by R. Levine, I. Chein, and G. Murphy (1), confirmed Sanford in showing that the food-need played a measurable role in the interpretation of a very vague, almost unstructured situation. The fourth, by H. Proshansky and G. Murphy (4), showed that reward and punishment associated with specific percepts could produce distortions in the direction of the rewarded percept, *e.g.*, *S*s who had received a money reward whenever a long line was shown them and who had been deprived of money as punishment whenever a short line was shown them, later revealed a significant tendency to overestimate the length of lines. The need for reward was effective in modifying perception; autism was active.

The Present Experiment

Purpose: Our purpose was to show the role of autism in the determination of a visual figure-ground relationship. Since money has been ingrained by conditioning as a means of fulfilling a wide variety of needs, the need to gain money was judged suitable for the establishment of the autism. The technique was to build up a strong association between certain visual patterns and reward, and between certain other visual patterns and punishments, later throwing a 'reward'

* An honors study conducted by Roy Schafer under the direction of Gardner Murphy.

pattern together with a 'punishment' pattern into an ambiguous situation so designed as to permit the perception of one pattern at the expense of the other.

Specifically, a reversible contour line was designed such that it could be seen as either of two distinct profiles looking in opposite directions, right and left. Eye-dots were added to help determine the faces. By making the profile the boundary of a vertical semicircle, either one of the two possible faces within the whole circle could be determined; an ambiguous situation was created in which either face might be seen as *figure* with the opposite aspect as ground. There were two such pairs of 'complementary' faces, shown in Fig. 1. These four faces were presented in random order one at a time; they were pairs from the experimenter's viewpoint, separate faces from the subject's viewpoint.

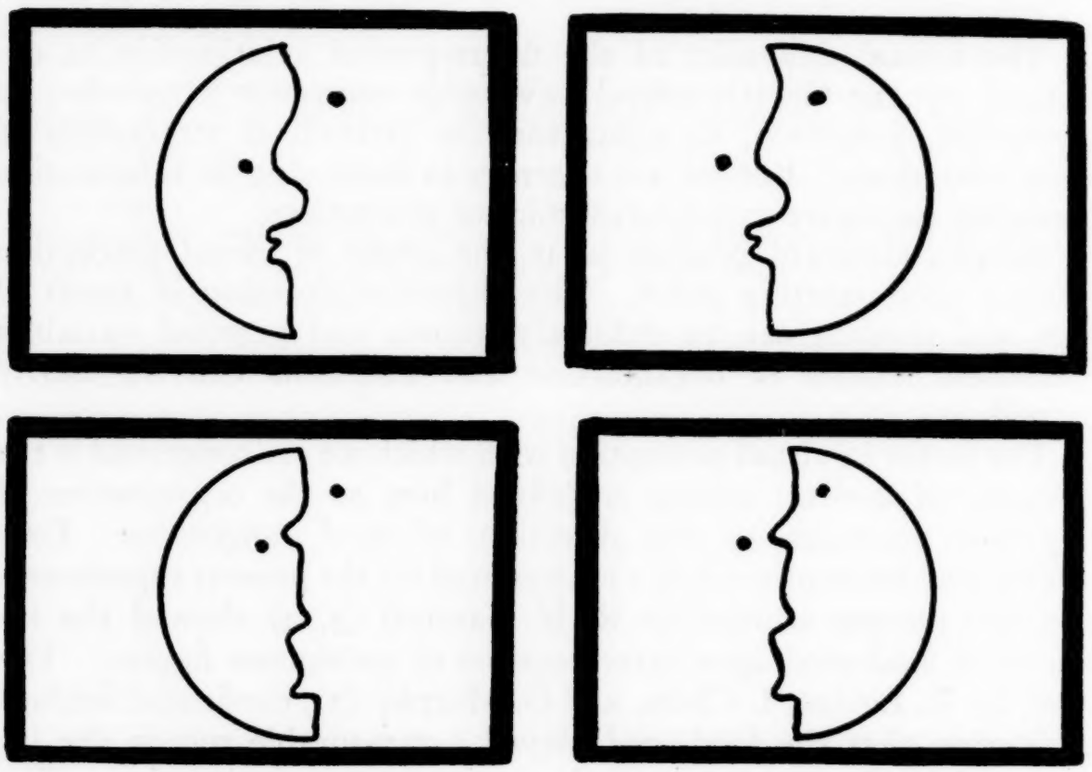

FIG. 1. The four stimuli presented in the training series.

To prevent the reversal of fixation during each presentation (which would make it possible for *S* to see both the complementary faces), the faces were presented on a Whipple tachistoscope for about one third of a second. This time was found to be optimal for permitting a single adequate fixation of the face.

The training period was conducted in the following manner: *S* was told that he would many times be shown on the tachistoscope each of four distinct faces. Before each presentation *E* would tell *S* the name he was to associate with the following face. In the course of the experiment *S* was to learn the four faces and their associated names. To *S* this was a *learning* experiment. He was told that whenever he saw either of *two* of the four faces he would be rewarded by two or four cents and that whenever he saw either of the *other two* faces he would be punished, by the loss of two or four cents. After *E* had told *S* orally the name of the face he was about to present, and after *S* had repeated that name aloud (to enforce the association), *E* told *S* how much he had won or lost. *S* was then to take from a pile of about fifty pennies before him the appropriate amount, or to return the appropriate amount from his own pile. The concreteness of this reward-punishment situation was felt advisable, to reinforce the effect of sheer knowledge of winning or losing.

A random order of 100 presentations was prepared, 25 for each of the four faces. The rewards and punishments were distributed by *E* among the two 'reward' faces and the two 'punishment' faces in advance, in such fashion that there were slightly more four-cent rewards than four-

cent punishments. This was done to permit *S* to build up gradually, through the hundred presentations, a comfortable balance, and thereby to prevent *S* from developing a negative attitude toward the whole experiment, including the rewards. We could expect such an attitude if *S* constantly hovered around the zero point and none of the reward money was retained. The number of reward experiences with each 'rewarded' face was, however, always 25; the number of punishment experiences with each 'punished' face was always 25. Each *S* won about 15 cents by the end of the experiment.

A second technique to reinforce the impact of rewards and punishments was introduced. *S* was told that the experiment would be run like a game of chance, *S* calling out a number from one to four before each presentation. *E* renumbered the four faces each time, so that for each presentation any number might be attached to any of the four faces. A slight hint was dropped that there might be a *system* in the distributions. This led all the *S*s to believe that they were

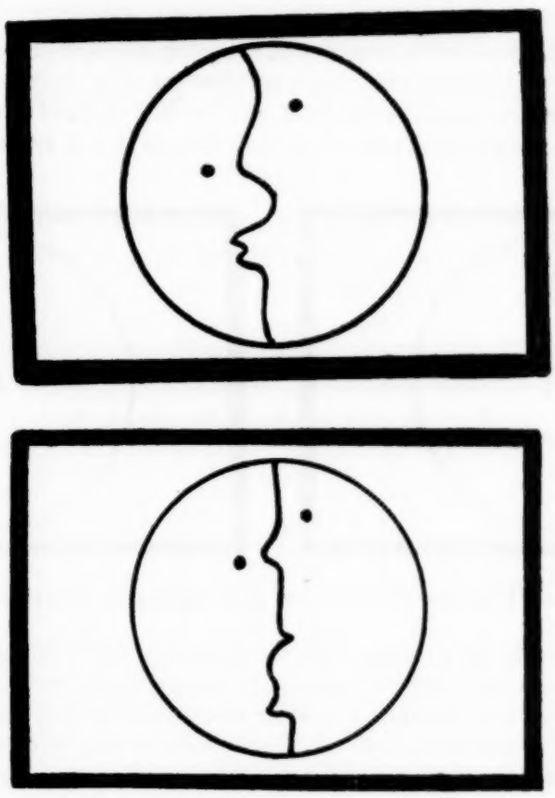

FIG. 2. The two stimuli presented in the post-training series.

picking the faces they saw and, indirectly, picking their rewards and punishments. This, as well as the tendency to try to figure out the system of picking all the reward faces, served to induce active anticipation of each perception and its attached reward. In this way interest and motivation were kept high throughout the experiment. Of course *S* played no part in determining the faces he really saw, since the series had been previously worked out. No *S* ever became aware that he was being deceived.

Several further precautions were introduced in the training period. One was to provide on the table before *S* rough copies of the faces to be learned, to which he was encouraged to refer before the first few perceptions of each of the faces. He was told that this was necessary if he was to perceive the faces as meaningful, since at first sight they were strange representations, especially when seen for only one-third of a second. Although such confusion was a real possibility, the main reason was to prevent *S* from perceiving the opposite aspect of the face he was supposed to perceive. (Such confusion had been observed to occur in preliminary subjects who did not enjoy the aid of the crude representations.) These 'aid' sketches were drawn to make the profile absolutely non-reversible, and no confusion was observed to arise in the present study.

A second precaution was used to control the 'dominance' factor for the four faces. That is, one member of each pair might in some way be dominant over its partner under *any* set of circumstances, and thus interfere with the action of the rewards and punishments. The following control was introduced: the first two *S*s were rewarded when they saw the right-pointed face of the first pair and the left-pointed face of the second pair. The next two *S*s were punished for the right-pointed face of the first pair and the left-pointed face of the second pair. Rewards were thus given to faces looking in opposite directions; and the same for punishments. The fifth *S* was rewarded for both of the left-pointing faces.

A third precaution was to threaten *S* after about 80 perceptions (20 perceptions of each face) that within the next two dozen presentations a 'phony' would be slipped into the series; if *S* missed it he would be penalized 10 cents. It was explained (truthfully) that this was to insure their paying attention to the faces. Actually one 'phony' was used and all the *S*s recognized it.

The training series was divided into two sets of 50 perceptions separated by a five-minute rest period.

The post-training critical series was presented after another five-minute rest period, during which *S* was given his instructions. He was informed that there would be no more rewards or punishments. He was to demonstrate how well he had learned the faces and their names. His task was merely to designate the name of each face as it was presented. He was informed

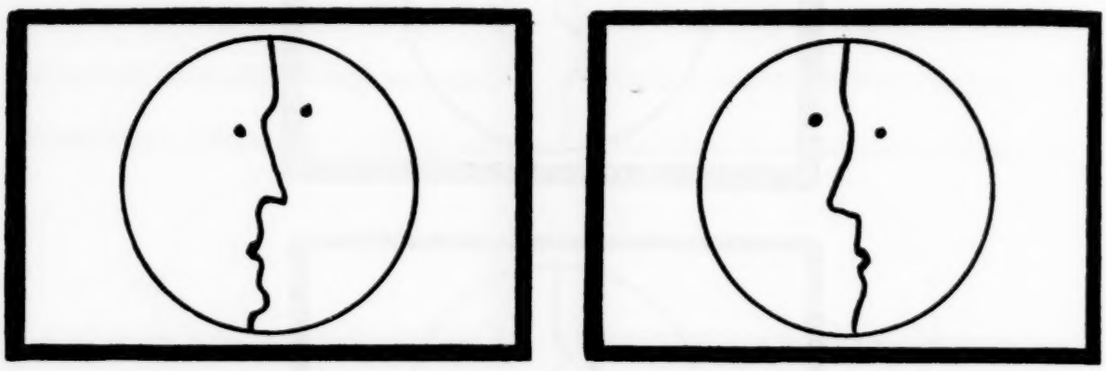

FIG. 3. The two stimuli used in breaking up direction sets.

that some of the faces would be strange, since we wished to see if he could distinguish old faces from new. All the faces would now be enclosed in complete circles (no explanation of this was offered.) *S* was to report aloud, as well as to write down on a sheet of paper before him, the name of each face, or an 'x' if it was unknown. The purpose was simply to ascertain whether, in the ambiguous situation, the previously rewarded faces predominated as figure in the figure-ground relationship.

We referred just above to new and old faces. The new faces, not appearing in the training series at all, are of importance in breaking *S*'s 'direction set.' These unknown faces were drawn and presented in the same fashion as the ambiguous faces already used; but each circle contained only one acceptable 'face,' a one-way profile line with a dot for the eye, and permitted reversal only with considerable effort. There were two such non-reversible patterns, one looking to the right and one looking to the left. As *S* reported aloud the face he had just seen on an ambiguous card, *E* presented for the next identification that strange face which was looking in the *opposite* direction from that reported. In preliminary experiments this had been found necessary, since *S*s tended to show a 'direction set' if ambiguous cards only were presented; they would see all the faces looking consistently to the right or consistently to the left. This set was so strong that when an ambiguous card was presented for which only one component face had been presented to *S* in the training period, *S* would not see that face if the 'direction set' favored the opposite direction. He would perceive the opposite face which he had never seen before and would therefore report it as an unknown face. The technique used here was designed to prevent a direction set by reversing the direction of each perception of an ambiguous card on the following card.

At least 32 presentations of the ambiguous cards of Fig. 2 (16 for each pair of composite faces) were given to each *S*.

Subjects: Six volunteer *S*s were used, four from introductory psychology classes and two who were entirely unfamiliar with psychology. The system of doling out rewards and punishments kept all the subjects highly motivated. None became aware of the reversibility of the contour line during the experiment proper.[1] One *S* was used only in a procedure described later.

Results

The first thing to note is that the perceptions of all five *S*s were predominantly in the same direction—the rewarded direction—for the first 16 presentations of the post-training series. Eight presentations of each ambiguous situation were given. The data from the two situations will be treated together for the most part. It can be seen that some *S*s show more consistency than others; this is best seen in Table II, which summarizes the date of Table I.

TABLE I

Sequence of the First 16 Perceptions for All Five *S*s

	Responses															
S	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	x	r	r	r	r	p	r	r	r	r	r	p	r	p	r	r
2	r	r	r	x	x	r	x	p	r	r	r	p	x	p	r	r
3	r	r	x	x	r	r	r	x	r	r	r	r	r	r	r	r
4	r	r	r	r	p	r	p	r	r	p	r	p	p	r	p	r
5	p	r	r	r	x	r	r	x	r	r	r	r	x	r	x	x

Note: r means that a rewarded face was perceived; p means that a punished face was perceived; x means that *S* failed to recognize either of the opposing aspects.

TABLE II

Summary of the First 16 Perceptions for All Five *S*s

S	R	P	X	N
1	12	3	1	16
2	9	3	4	16
3	13	0	3	16
4	10	6	0	16
5	10	1	5	16

[1] Naiveté is, however, probably not essential. By the present method one might use as data the *first* face seen by *S* in an ambiguous situation provided that *S*, at a conscious level, honestly controlled his perceptions so as to favor neither of the two opposing aspects.

Table III shows the breakdown of responses for each of the two ambiguous situations utilized. It can be seen that 12 of the 13*x*'s occurred for the second ambiguous situation. This might be due to the greater acceptability of the first situation as really offering a

TABLE III

CLASSIFICATION OF PERCEPTIONS FOR THE TWO AMBIGUOUS SITUATIONS

Sit.	R	P	X	N	R%	P%
No. 1	32	7	1	40	82	18
No. 2	22	6	12	40	79	21

Note: the base for the percentages was N − X.

'face,' or the greater fluidity, or ease of fixation-reversal, in the second situation. Such fluidity might be a factor making for confusion. The percentage of R's and P's for the two situations is, however, about the same.

There are several ways in which the data may be statistically treated. It seems to us that since the *x*-response is simply a failure to recognize a face which has been repeatedly shown, the most reasonable procedure is to exclude the *x*-response, and deal simply with the recognized faces, asking ourselves simply whether in the face of such recognition, reward and punishment make a difference. From this viewpoint, it may be seen from Table IV that 54 of the

TABLE IV

SIGNIFICANCE OF THE R − P DIFFERENCE BY THE CRITICAL RATIO METHOD

R	P	N	St. Error	R − N/2	C.R.
54	13	67	4.1	20.5	5

total meaningful perceptions of ambiguous situations were perceptions of faces to which rewards had been attached in the training period. Only 13 perceptions were of punished faces, as many as were left unidentified. A chance proportion among the R + P total would have been R = N/2 and P = N/2. The standard error of a chance distribution of 67 such responses is 4.1 perceptions. The difference of the number of R-perceptions from that expected by chance (namely 33.5) is 20.5. 20.5/4.1 gives us a critical ratio of 5.

Other methods of treating the data may be preferable, but with such a large margin of safety, it seems hardly worth while to discuss their various merits.

Discussion

Although at least 32 ambiguous situations were presented to each *S*, we have used only the responses to the first sixteen presentations. It was observed that a new kind of set was established around the sixteenth perception which made all further perceptions irrelevant to our hypothesis. This set could apparently function independently of previous reward and punishment experiences. If we think in terms of conditioning we can see that the conditioning for or against certain faces would break down if the rewards and punishments were removed. It apparently took about 8 experiences with each ambiguous situation to break down the conditioning which we had set up.

The set may be described as one which caused *S* consistently to see one or the other of the two faces in each ambiguous situation, so that in the case of 4 of the 5 *S*s, from the sixteenth perception onwards only one face in each pair was perceived at all. The fifth *S* established his set after about 50 presentations. It was as if *S* had come to react in a global manner to the general situation of a curved or angular contour line within a circle by identifying that situation with one of the faces it contained. Fineness of perception was no longer necessary, and *S* was secure with his consistent interpretation (which was a *correct* interpretation after all). From the 16th to the 32nd perception this set was undoubtedly reinforced (it was a swift, correct, satisfying response) for those four subjects who consistently held to their sets. These sets are sometimes in the *same* direction, but with about equal frequency in the direction *opposite* to, that of the earlier reward-punishment effect.

That there was strong reinforcement is supported by further evidence obtained from the *S*s. *E* tried to break down the sets after 32 presentations. First *E* asked *S* to 'image' as clearly as possible one of the faces suppressed by the set. *S* described the face to *E* quite accurately in every case; there was no *forgetting*. Then *S* was told that the next face he would see would be the one he was imaging. The ambiguous card containing that face was then presented, and *S* was asked which face he had seen. In no case did this technique break down the set at once; it was effective on some of the *S*s after several repeats. In other words, *S* did not see the face he had imaged if it ran against the set he had established between the 16th and 32nd presentations. He failed to realize that what he had imaged was the reverse, or complement, of what he now saw.

A second test of the stability of the set was used. *S* was shown the suppressed faces on the original cards, where they were clearly determined, at least twice in succession. All the *S*s recognized the faces under these circumstances. However, when they were told

that they would see next the same face enclosed in a circle, and were then shown the corresponding ambiguous card, they failed to see that face and saw instead the opposite aspect, reporting this to *E*. Two previous clear-cut experiences with a face were ineffective in causing it to be perceived in the ambiguous situation if there was a set to structure the ambiguous situation in a different way. *After a short period during which the induced autism was active, the perceptual processes underwent a process of consolidation or fixation, depending on factors within the organism other than the earlier reward-punishment experiences.*

A control *S* was used. He received no rewards or punishments but was otherwise treated the same way as the experimental subjects. He too acquired a set for two particular faces in the post-training series after about 16 presentations. This phenomenon may therefore be seen as acting independently of rewards and punishments, and one worthy of study in itself.

Such observations as these lead to a preliminary hypothesis to the effect that in a constantly ambiguous situation an individual will in time fixate a particular way of structuring the situation so strongly that all ambiguity value will be lost; the individual will perceive the situation as a clearly defined one. This concept is most relevant to the dynamics of the fixation of stereotypes and norms. One might expect individual differences in the rate of acquiring such fixations, and in the strength with which they will be retained. The fifth *S* achieved his fixations or sets only after about 50 presentations and abandoned these sets more easily under experimental pressure than any of the others. This interesting by-product of the experiment needs further investigation.

An attempt to follow up the relation between the ultimate fixations and affective ratings of the faces by three of the *S*s several weeks after the experiment proper showed that two of the *S*s rated their final fixations above the opposite aspects of these fixations while a third *S* rated that face which he had most strongly fixated, and which had been a reward face, lowest. Nothing reliable can be said about the affect attached to the fixations on the basis of these few inconsistent cases.

One improvement in the testing for the presence of the autism may be suggested; the faces, both rewarded and punished, should be presented on reverse sides of the contour line, as they would be seen *from the other side of the original card*, in the ambiguous situation. This would be a further check on whether *S* needs to see rewarded faces only. Factors of position and direction should be ineffective if the autism is strongly at work.

Summary and Conclusions

This experiment sought to trace the role of autism in the determination of figure-ground relationships in visual situations. Opposite complementary aspects of contour lines—each being capable of being seen as a face—were presented separately to 5 *Ss* on the Whipple tachistoscope. Each one of the aspects had been strongly defined on a separate card by making it the boundary of a vertical semicircle. Rewards were given in association with certain of these faces, and punishments in association with the opposing aspects of these rewarded faces. Reward or punishment, respectively, was given 25 times for each of the component faces to build up a strong association. Then the contour line was presented in a complete circle, in which situation either of the opposing aspects of the contour line could be seen. Sixty-seven perceptions of faces were gathered, in all, from the 5 *Ss*; 54 of these (a significantly high number) were perceptions of faces which had been rewarded in the training series. The main hypothesis has been verified: autism can function in the determination of the figure-ground relationship in a visual field.

An interesting by-product emerged. After discontinuance of the reward-punishment procedure, each *S* showed a consistent tendency to consolidate his perceptual pattern so as to see only one, never the other, aspect of a pair, and this consolidation tendency appeared to be independent of the earlier rewards or punishments. This consolidation tendency merits further analysis.

(Manuscript received October 26, 1942)

References

1. Levine, R., Chein, I. & Murphy G. The relation of the intensity of a need to the amount of perceptual distortion: a preliminary report. *J. Psychol.*, 1942, 13, 283–293.
2. Proshansky, H. & Murphy, G. The effects of reward and punishment on perception. *J. Psychol.*, 1942, 13, 295–305.
3. Sanford, R. N. The effects of abstinence from food upon imaginal processes. *J. Psychol.*, 1936, 2, 129–136.
4. Sanford, R. N. The effects of abstinence from food upon imaginal processes: a further experiment. *J. Psychol.*, 1937, 3, 145–159.

STUDIES IN AUDITORY THEORY

II. The Distribution of Distortion in the Inner Ear *

BY WILLARD R. THURLOW

Princeton University

Distortion, in audition, consists in the appearance of overtones in the response to a single pure tone stimulus, and of combination tones in the response to two pure-tone stimuli. Recent experiments (**6–10, 1**), making use of the electrical responses of the ear, have shown that the major pattern of distortion arises in the inner ear.

A question of great importance for auditory theory is whether the pattern of disturbance in the inner ear caused by a given overtone or combination tone is the same as that caused by a tone of corresponding frequency reaching the inner ear from an external source. To take a specific example: Suppose the ear to be stimulated by a strong pure tone of 1000 cycles; the first overtone will be present in the response of the inner ear, and will have a pitch corresponding to that of a 2000 cycle tone impressed on the ear from an external source. The subjective pitch is the same in the two cases; according to a 'place' theory the disturbances in the inner ear should have the same distribution. If the two distributions should be found to differ significantly, a 'place' theory of hearing, which postulates that pitch is determined by the spatial character of the disturbance on the basilar membrane, would be shown to be inadequate.

It was the object of the present experiment to compare these two distributions for the case of overtones.

Method

Cochlear potentials were utilized as a measure of inner ear activity, and were recorded from the cochlea of the guinea pig. The guinea pig was chosen as the experimental animal since the cochlea in this animal extends almost completely into the middle ear, thus making it possible to record from a series of points. The method consisted essentially in equating the response of the overtone and its corresponding frequency at one point on the cochlea; then of comparing the responses at each of several other points on the cochlea, using the same intensities of stimulation as had been used at the first recording point.

The experimental animal was anaesthetized with urethane. The bulla acoustica was exposed by a ventral operation through the neck; removal of a portion of the bulla left the cochlea beneath in a favorable position for recording. The active electrode, consisting of a

* From the Psychological Laboratory of Princeton University. The author would like to express his appreciation to Dr. Wever and Dr. Bray for the use of the research equipment with which the present experiment was performed. He would also like to thank them for their constant interest and suggestive criticism during the course of both this and the preceding experiment.

copper wire, was placed successively on each of the seven positions indicated in Fig. 1; the ground electrode was placed on the muscles of the neck. The cochlear potentials resulting from tonal stimulation were led from the electrodes through an amplifier to a wave analyzer (General Radio, Type 736A). The tonal stimuli were produced by a system consisting of oscillator, filter, at-

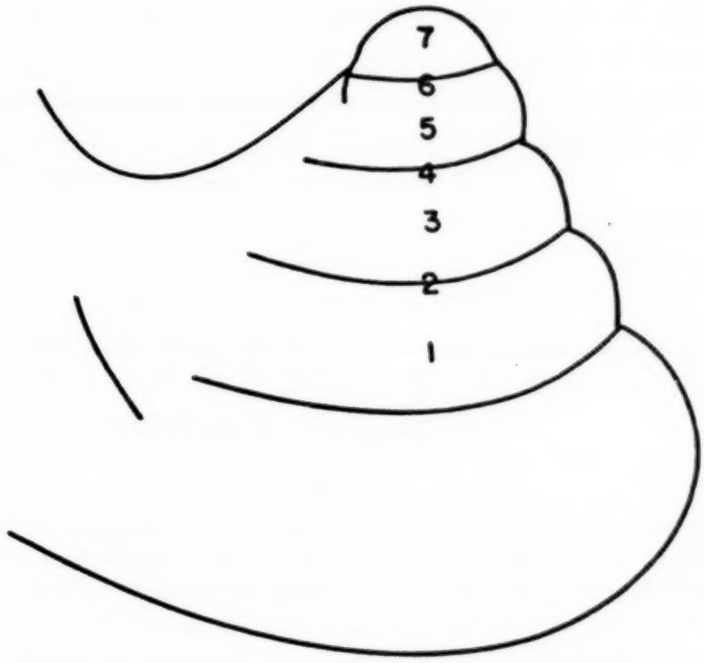

FIG. 1

The cochlea of the guinea pig is represented diagrammatically, with the recording positions numbered 1 through 7, a progression from basal to apical turns.

tenuator, amplifier, and speaker. The tones were led to the ear of the animal from the speaker by means of a closed tube.

The specific recording procedure was as follows: With the active electrode on position 1, the response of the overtone was adjusted by manipulating the intensity of the fundamental so that a response of approximately one microvolt was obtained. (This was near to the maximum response obtainable.) The intensity in decibels was noted, and this tonal stimulus was then turned off. The frequency corresponding to the overtone was next introduced at an intensity which gave an amount of response equal to that which the overtone had given. This intensity was also noted. Responses were then taken from the remaining points for both the overtone and its corresponding frequency, using the same intensities as had been used at the first recording point.[1] Finally, responses were taken at the original recording point, in order to be certain that the responses were still equal at this point.

Observations were made on a total of 10 guinea pigs, observations on the first two being of a preliminary exploratory nature.

RESULTS

Data obtained with the second overtone of 1000 cycles, corresponding to a frequency of 3000 cycles, are shown in Fig. 2, for guinea pig Number 3. The ratio of overtone response (R_O) in microvolts to the response to the equivalent 3000 cycle tone (R_E) is plotted

[1] In several of the experiments, position 1 was omitted because of the small amount of response obtained. This is due to the thickness of the bone at this point. If position 1 was omitted, position 2 was used as the first recording point.

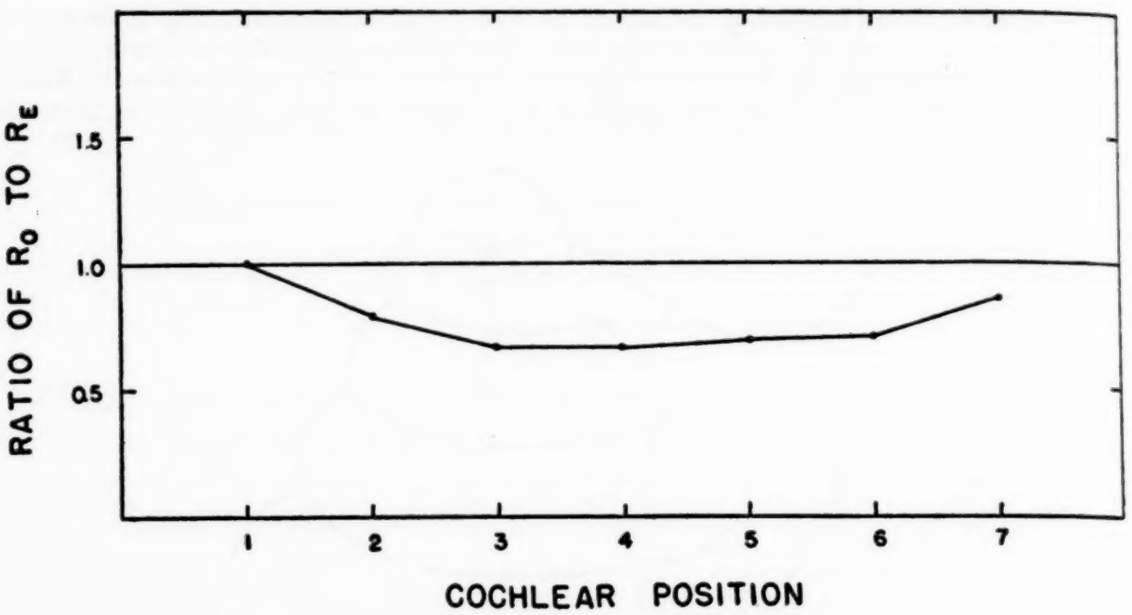

FIG. 2

Data obtained with the second overtone of 1000 cycles, corresponding to a frequency of 3000 cycles, are shown. The ratio of overtone response (R_O) in microvolts to the response to the equivalent 3000 cycle tone (R_E) is plotted as a function of cochlear position.

as a function of cochlear position.[2] In progressing from position 1 to position 7, the response to both overtone and corresponding frequency was found in general to increase;[3] but the response to the overtone increased to a much smaller extent. Thus the ratio tends to decrease in going from position 1 to position 7. A recheck at the original position showed the responses at this point to be still equal. Similar results were obtained with guinea pig Number 5. Results on guinea pig Number 4 were variable; there was a difference at only two points, and at these points the overtone gave a greater response than the 3000 cycle equivalent tone. The difference in the results with this animal may have been due to the collection of fluids in the middle ear around the cochlea; such a collection might possibly so alter impedance relations, that if there was a difference in distributions, the relations of recorded response might be reversed from that ordinarily found.

In order to check the possibility that chance variation affected the results, determinations of the reliability of the difference found at one of the recording positions were made with subsequent animals, for each of the overtones used. At the same time, the extent to which the responses differed at the original equating position in a

[2] Ratios rather than differences are plotted, since the factor by which response is reduced due to impedance between place of generation of the potentials and place of recording will not affect a ratio, as it would absolute difference values.

[3] This increase in response is probably due to the decreased thickness of bone, and corresponding decreased impedance. The same result has been obtained by other experimenters.

number of readings was determined, thus providing a more exact check on the degree to which changes in the animal's condition might be operating to influence the results.[4] For the readings at each of the two points, determinations were made of the significance level at which the mean difference between the responses to the overtone and the responses to the equivalent frequency was different from zero. (The ratio of the mean difference to the standard error of the mean difference was determined; reference was then made to a table of the *t* distribution (**3**)).

This was done for the second overtone of 1000 cycles, equivalent to 3000 cycles, in guinea pig Number 7, at positions 2 (the 'equating' position) and 6. At position 6, the response to the overtone was significantly smaller than the response to the equivalent 3000 cycle tone beyond the .01 level. The responses were not significantly different at position 2.

In Fig. 3 are shown the results for the first overtones of 1500, 1000, and 400 cycles, corresponding to frequencies of 3000, 2000, and 800 cycles respectively. For these overtones position 2 was used as the 'equating' position, and position 1 was omitted, since the response from position 1 was too small. The data for the 3000 cycle overtone were obtained from guinea pig Number 9. The responses of overtone and equivalent frequency were significantly different at point 6 beyond the .01 level. Readings at point 2 were not significantly different from each other. Results on reliability identical with this were obtained with guinea pig Number 8. The overtone response was similarly smaller at point 6.

The data for the 2000 cycle overtone were obtained with guinea pig Number 9. Differences significant beyond the .01 level were found at position 6 with both this animal and guinea pig Number 7, the overtone response being smaller in both cases. No significant differences were found at point 2.

The data for the 800 cycle overtone were obtained from guinea pig Number 10. The overtone response at point number 6 was significantly smaller than that to the equivalent 800 cycle tone beyond the .01 level; and the readings of each, taken at point 2, were not significantly different from each other. As shown in Fig. 3, the overtone response fell near the apex to 35 percent of the response to the equivalent tone of 800 cycle frequency. For purposes of comparison, the responses to 4000 cycles and 400 cycles were equated at point 2 (guinea pig Number 6), then compared at point 6, and finally re-

[4] In more detail, the procedure was the following: At the original recording position, after the overtone and equivalent frequency had been equated, 6 readings were taken of each; 6 readings of each were then taken at a point where a difference was found; finally 6 more readings of each were taken at the original recording position.

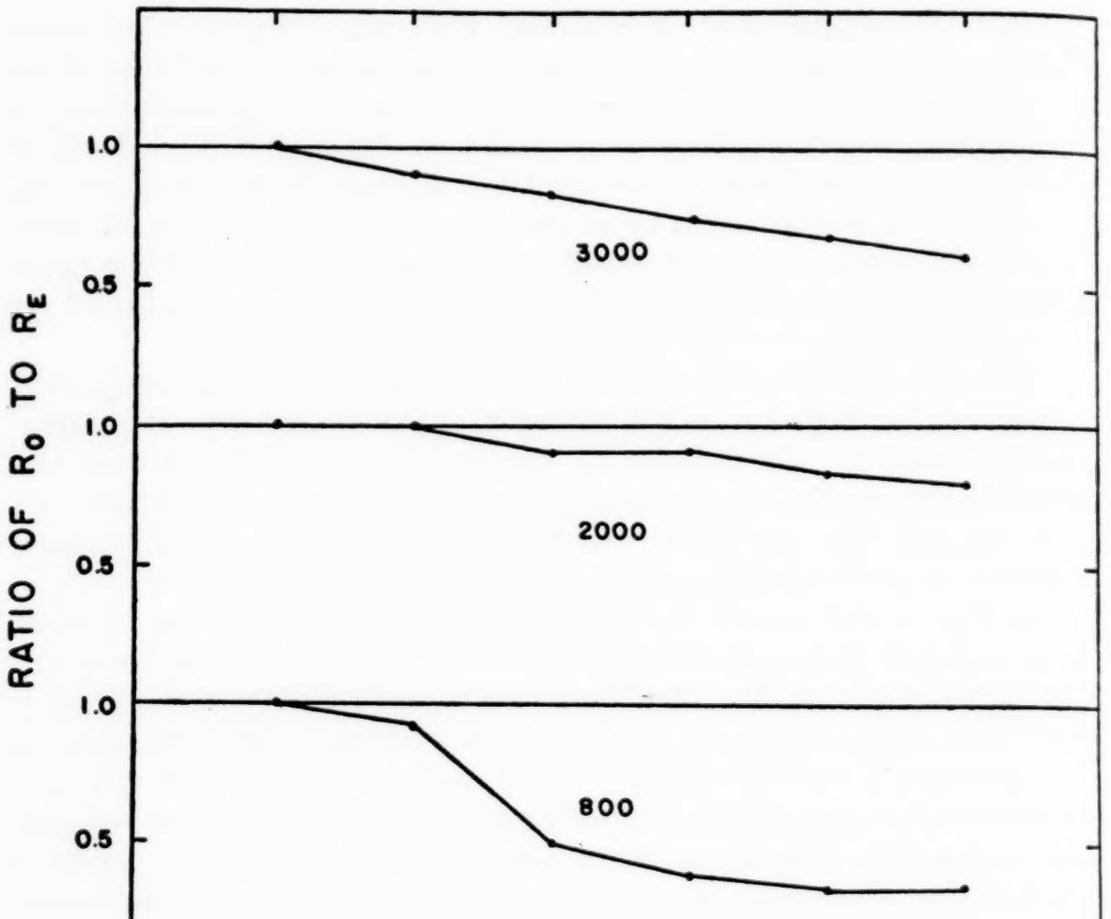

FIG. 3

The results for the first overtones of 1500, 1000, and 400 cycles, corresponding to frequencies of 3000, 2000, and 800 cycles respectively, are presented. As in Fig. 2, the ratio of overtone response (R_O) in microvolts to the response to the equivalent frequency (R_E) is plotted as a function of cochlear position.

checked at point 2. At point 6, the response to 4000 cycles was 40 percent of that to 400 cycles. Of course, the comparison of these percentage values is merely suggestive; for the impedance between position of generation of the potential and recording position changes in going from position 2 to position 6, and undoubtedly does not have the same effect upon the response to the 400 cycle tone as upon the 4000 cycle tone because of the differing frequencies.[5]

[5] In comparing the distribution of response within the cochlea for different frequencies, Johnson (4) has attempted to correct for impedance differences by subtracting the same corrective factor at a given recording point for all frequencies. Culler (2) has not attempted to correct for the effects of different impedance. The results of these experiments are therefore difficult to interpret. (A further obstacle in the way of interpretation of these same experiments has been

Discussion and Conclusions

The foregoing experiments have shown that although an overtone and the frequency from an external source corresponding to it have the same pitch, they do not have the same distribution in the inner ear. It should be possible to extend the method of analysis used in the present experiment to the case of bone conducted tones. Thus, an air-conducted and bone-conducted tone of the same frequency have the same pitch; test could be made to see whether the distribution of their disturbances in the inner ear is the same. Test could also be made by this method to determine whether the form of the distribution of disturbance in the inner ear caused by a given tone remains the same when another tone or tones are introduced.

The results of the present experiment indicate that pitch is not mediated by the spatial form of the disturbance which occurs in the inner ear as a result of tonal stimulation. It is justifiable also to conclude that pitch is not determined centrally by the spatial character of the disturbance. It should be possible to gain evidence on the validity of this latter statement by the application, at higher levels of the auditory system, of the method of analysis used in the present experiment.

In a previous experiment (**5**) it was shown that a change in the pitch of one tone could be produced by introducing a tone in the opposite ear, provided both tones were of high intensity. It was furthermore proved that the change in pitch was not due to a peripheral interaction; that the central interaction was not of a type which could be explained in spatial terms.[6] The results of the present experiment are in agreement with those of the previous experiment in indicating the inadequacy of an explanation of the mediation of pitch by position of maximal stimulation, or other spatial characteristic.

Summary

The distribution of the disturbance of an overtone in the inner ear (produced by stimulating the ear with a loud pure tone) is found to be significantly different from that of a tone of frequency corresponding to the overtone, although both have the same subjective

the use by the experimenters of the number of decibels intensity necessary to produce a constant response as a measure of sensitivity. Thus, in comparing the results for different frequencies, the implicit assumption is made that the relation between cochlear response and intensity is linear, with equal slopes for each frequency. However, the intensity functions obtained when recording from the bony wall of the cochlea are often of unequal slope and sometimes of irregular slope. This latter statement is based on unpublished observations by the author, contained in Senior thesis, Brown University, 1939.)

[6] The possibility of an essentially non-spatial effect resulting in a secondary spatial effect was considered improbable.

pitch. It is concluded that the pitch of a tone is not correlated with a spatial characteristic of the disturbance caused by that tone in the inner ear; and that it is not probable that a spatial characteristic arises higher in the auditory system which is correlated with pitch. These conclusions agree with those reached in a previous experiment. Further applications of the method of analysis used in the present experiment are discussed.

(Manuscript received November 14, 1942)

References

1. Bray, C. W. & Thurlow, W. R. Interference and distortion in the cochlear responses of the pigeon. *J. comp. Psychol.*, 1942, **33**, 279–289.
2. Culler, E. In: Symposium on tone localization in the cochlea. *Ann. Otol., etc., St. Louis*, 1935, **44**, 809–815.
3. Fisher, R. A. *Statistical methods for research workers.* (6th ed.) Edinburgh and London: Oliver and Boyd, 1936, p. 166.
4. Johnson, E. P. Localization of response in the cochlea as determined by electrical recording. M.Sc. thesis: 1940 (filed in Brown University library).
5. Thurlow, W. R. Studies in auditory theory. I. Binaural interaction and the perception of pitch. *J. exp. Psychol.*, 1943, **32**, 17–36.
6. Wever, E. G., & Bray, C. W. Distortion in the ear as shown by the electrical responses of the cochlea. *J. acoust. Soc. Amer.*, 1938, **9**, 227–233.
7. Wever, E. G., Bray, C. W., & Laurence, M. The locus of distortion in the ear. *J. acoust. Soc. Amer.*, 1940, **11**, 427–433.
8. Wever, E. G., Bray, C. W., & Laurence, M. The origin of combination tones. *J. exp. Psychol.*, 1940, **27**, 217–226.
9. Wever, E. G., Bray, C. W., & Laurence, M. The interference of tones in the cochlea. *J. acoust. Soc. Amer.*, 1940, **12**, 268–280.
10. Wever, E. G., Bray, C. W., & Laurence, M. The effect of middle ear pressure upon distortion. *J. acoust. Soc. Amer.*, 1941, **13**, 182–187.

PROBLEM SOLUTION BY MONKEYS FOLLOWING BILATERAL REMOVAL OF THE PREFRONTAL AREAS. I. THE DISCRIMINATION AND DISCRIMINATION-REVERSAL PROBLEMS

BY H. F. HARLOW AND J. DAGNON *

University of Wisconsin

I. Purpose

The purpose of the following investigation was to compare the behavior of monkeys following bilateral prefrontal lobectomy with that of normal monkeys in the solution of a series of discrimination and discrimination-reversal problems (**3**).

II. Subjects

Four operated animals were used in this investigation. Two of these, numbers 54 and 55 were rhesus monkeys. One of the others, G5, was a green monkey, and the remaining animal, number 31, was a sphinx baboon. Six normal rhesus monkeys, numbers 21, 22, 50, 51, 52 and 53 formed a comparison group.

III. Apparatus

The apparatus used in these experiments has already been adequately described (**2**, **4**). Its essential characteristics are illustrated in Fig. I. The stimulus tray used in these studies is essentially the same as that devised by Klüver (**5**, p. 387) and described as "a modified 'form-board.'"

A set of 5 pairs of discrimination objects were used in preliminary training tests and included both planometric and stereometric objects. The stimulus-objects used in the regular tests were pairs of stereometric objects differing in many characteristics, and a typical pair of these is shown in Fig. I. A detailed description of the objects used in the preliminary training tests, and the reasons for their selection has been given in an earlier publication (**3**).

IV. Method

Operative procedures:

In all cases the bilateral prefrontal lobectomies were carried out in a single operation, employing the usual aseptic precautions. On the lateral surface of the cortex of both hemispheres the line of incision passed about one-half centimeter posterior to the terminations of both the superior and inferior limbs of the arcuate sulcus, approximately parallel with the central sulcus. After passing through the grey matter on the lateral cortical surface, the knife was moved downward and diagonally forward about a half centimeter, to the floor of the cranial vault and medially to the horizontal fissure. The tissue was then removed en bloc. These procedures insured complete ablation of the frontal association areas (Brodmann's areas 9–10–11–12) even though some damage to areas 6 and 8[1] resulted.

* This work was supported in party by a grant from the Special Research Fund of the University of Wisconsin for 1941–42.

[1] The operations on subjects G5 and 31 were made by J. A. Bromer and on 54 and 55 by H. F. Harlow.

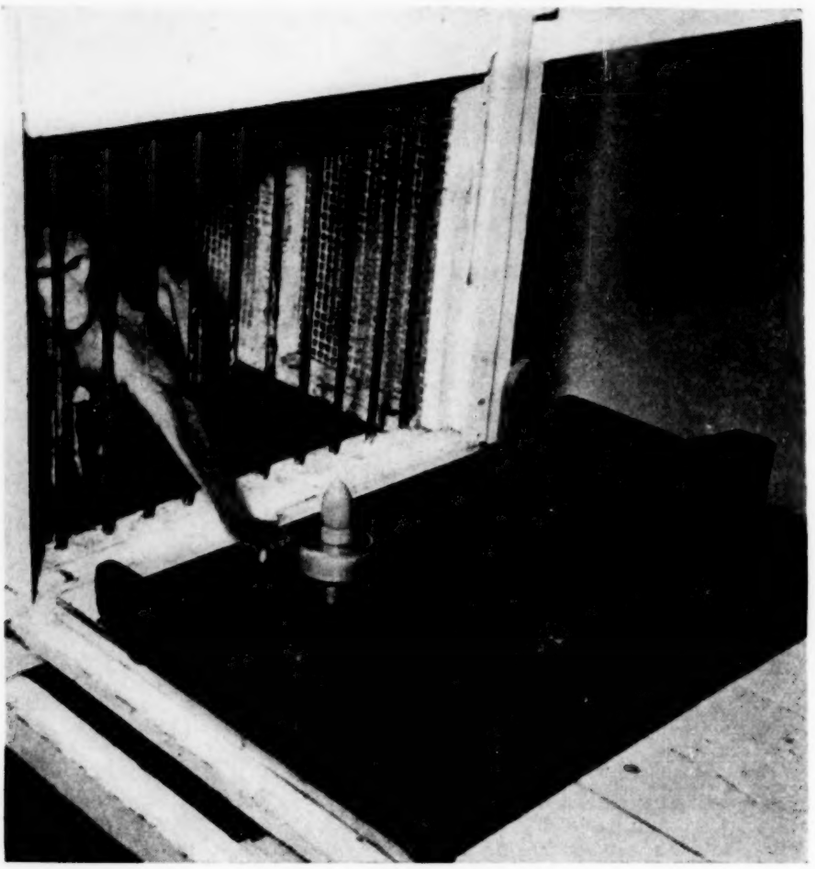

FIG. I. Solution of Discrimination Problem by Subject 54.

Taming and pre-test procedures:

Subjects G5 and 31 were not run on the present tests until over a year had elapsed after the operation. During this time they had been tested on a series of oddity and oddity generalization tests and had proved to be apt subjects.

Animals 54 and 55 were given about six weeks to insure recovery from the operation and were then tamed over a three months' period. In spite of constant and painstaking efforts made during this interval, the monkeys failed to adapt to the taming procedures with the success usually obtained on normal animals. Persistent and successful attempts by the subjects to bite the experimenter continued for weeks and are still present in somewhat sublimated form in subject 55 two years after the operation. Despite these distractions, taming, involving handling and petting the monkeys, was continued. Before testing was begun, special care was taken for a period of a month to adapt these operated animals to the general experimental situation. All in all, a period of nearly six months intervened between the operations and the beginning of the regular test procedures. Every effort was made to overcome the persistent lack of 'emotional restraint' indicated in these prefrontal subjects by difficulty of taming, restless activity, easy distractibility to external stimuli, and excited behavior in the presence of an experimenter to whom they had had every opportunity to become thoroughly accustomed.[2]

Aside from the *special care* taken in taming and in adapting the prefrontal animals to the test situation, subjects 54 and 55 were comparable to the six control animals in that they had had no previous test experience.

Test procedures:

The method was the same as that previously described by Harlow (3). Discrimination was between two stimulus objects placed over food wells on a tray which was pushed within reach of

[2] We have no data indicating that the lack of 'emotional restraint' was the result of the operation since the preoperative emotionality of the subjects was not tested. Furthermore, it should be noted that the operations involved more than the prefrontal areas.

the subject. Correct responses were rewarded with food. Following an incorrect response the tray was withdrawn, and the food shown to the monkey, but he was not allowed to obtain it.

On the preliminary tests 25 trials a day were run for each subject on one of the 5 preliminary discrimination objects until a criterion of 23 correct responses in the 25 trials of a daily run was attained.

In the *discrimination learning* tests each subject was tested on 15 pairs of stimulus-objects. A single discrimination was completed each day and the criterion of learning for each discrimination was 19 successive correct responses.

The discrimination training was followed by training for *discrimination reversal.* The monkeys were given re-learning trials on each of the first 10 pairs of discrimination objects used in the discrimination learning. As soon as any single discrimination was relearned (10 out of 10 correct responses) the discrimination was reversed, *i.e.*, the stimulus-object previously correct became incorrect, the stimulus-object previously incorrect became correct. Each reversed discrimination was run until the same criterion of 10 successive correct responses was attained. Thus, each day a discrimination was re-learned and its reversal then run until the criterion of learning was again reached.

V. Results

The number of errors made by the 6 normal and 4 operated monkeys in the solution of the 15 discrimination learning tests are shown in Table I. The mean number of errors made by the normal

TABLE I

Number of Errors Made by 6 Normal and 4 Prefrontal Monkeys in Solution of Each of 15 Discrimination Learning Problems

Normal Monkeys	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Mean
50	0	1	0	1	0	1	3	3	1	2	0	0	1	3	1	1.1
51	1	0	0	0	0	5	1	2	1	1	2	0	0	2	0	1.0
52	0	1	0	2	2	0	0	2	1	1	2	0	0	0	0	0.7
53	1	4	0	1	0	2	1	0	0	3	3	0	0	1	7	1.5
21	2	4	2	3	5	2	0	1	1	0	0	2	1	0	0	1.5
22	2	2	1	6	1	5	4	0	0	0	2	1	2	3	1	2.0
Mean	1.0	2.0	1.5	2.2	1.3	2.5	1.5	1.3	0.7	1.2	1.5	0.5	0.7	1.5	1.5	1.3
Operated Monkeys																
54	0	5	0	0	2	8	1	0	1	5	1	0	1	1	2	1.8
55	0	1	1	3	2	12	5	1	1	6	4	1	4	1	1	2.9
Jr	0	7	0	0	0	5	2	1	3	1	3	1	4	0	7	2.3
G5	1	1	7	7	1	10	5	8	2	7	10	4	3	0	2	4.5
Mean	0.3	3.5	2.3	2.5	1.3	8.8	3.3	2.5	1.8	4.8	4.5	1.5	3.0	0.5	3.5	2.9

monkeys is 1.3, as compared with the mean of 2.9 errors made per discrimination problem by the prefrontal subjects. Fischer's t value (1) for the difference between the operated and unoperated animals is 3.08. The difference is therefore significant at the 2 percent confidence level. If only the operated rhesus, 54 and 55, are compared with the normal animals, the difference is significant at the 5 percent confidence level.

Fig. II presents in graphic form the cumulative percentage of the 15 discrimination problems solved at various error frequencies by the two groups of subjects. Though the normal monkeys were consistently superior, the operated monkeys solved almost half

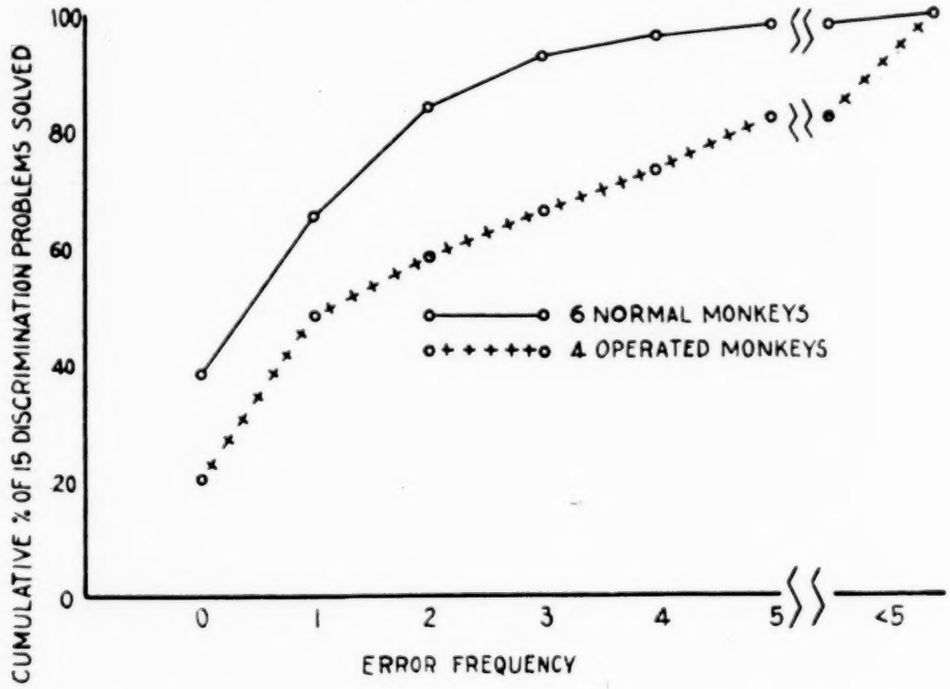

FIG. II. Cumulative Percentage of 15 Discriminations Solved at Various Error Frequencies.

the discriminations with not more than one error, and solved over half the discriminations after making an error score of two or less.

The number of errors made by the normal and operated monkeys in the 10 *discrimination-reversal* tests is presented in Table II. The mean number of errors made by the normal monkeys is 2.1 compared with the mean of 8.7 errors made per discrimination reversal problem by the operated animals. The *t* value for the difference between the operated and unoperated animals is 2.83. This difference is significant at the 5 percent confidence level. If only the operated rhesus, 54 and 55, are compared with the normal animals, the difference is significant at the 5 percent confidence level.

Fig. III presents in graphic form the cumulative percentage of the 10 discrimination-reversal problems solved at various error frequencies by the two groups of monkeys. The normal group is consistently superior to the operated group.

The difference between discrimination reversal and discrimination is greater in the case of the operated animals as indicated by a *t* value of 2.54, based on 8 degrees of freedom. If only the operated rhesus, 54 and 55, are compared with the normal animals, the difference is significant at the 20 percent confidence level.

TABLE II

NUMBER OF ERRORS MADE BY 6 NORMAL AND 4 PREFRONTAL MONKEYS IN THE SOLUTION OF EACH OF 10 DISCRIMINATION-REVERSAL PROBLEMS

Normal Monkeys	Discrimination Reversal Number										
	1	2	3	4	5	6	7	8	9	10	Mean
50	4	5	1	6	5	0	0	1	0	1	2.3
51	2	2	1	3	1	1	2	0	0	0	1.2
52	25	2	1	1	2	11	0	0	0	0	4.2
53	3	4	0	2	0	6	1	2	1	4	2.4
21	1	2	0	0	3	2	1	1	2	0	1.2
22	2	1	2	1	4	1	0	0	0	0	1.1
Mean	6.2	2.7	0.9	2.2	2.2	3.5	0.7	0.7	0.5	0.8	2.1
Operated Monkeys											
54	11	7	2	6	1	7	1	0	1	1	3.7
55	4	9	5	24	4	5	4	7	5	3	7.0
Jr	12	11	7	9	8	3	7	8	5	3	7.3
G5	8	41	14	22	16	14	20	20	6	8	16.9
Mean	9.0	17.0	7.0	15.3	7.3	7.3	8.0	9.0	4.0	3.8	8.7

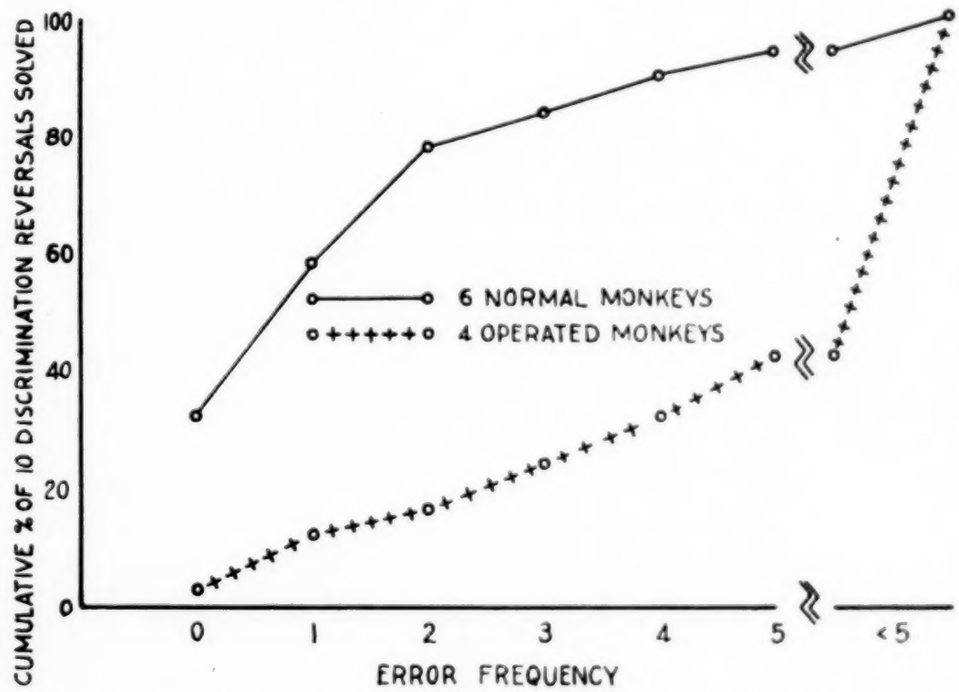

FIG. III. Cumulative Percentage of 10 Discrimination Reversal Problems Solved at Various Error Frequencies.

VI. DISCUSSION

The data of the present experiment show that monkeys following bilateral prefrontal lobectomy make more errors than normal monkeys in learning both a series of discrimination problems and a

series of discrimination-reversals, and the differences in number of errors made by the two groups approach statistical significance in both tests.

The differences do not result from a failure on the part of the operated animals to learn either discriminations or discrimination-reversals with reasonable facility. Thus the operated animals averaged less than 3 errors in solving the discrimination problems and less than 9 errors in solving the discrimination-reversal problems.[3]

Overlapping between the operated and normal groups appeared in that subject 54 made a lower average error score on the group of discrimination problems than the worst average error score by a normal monkey (subject 22), and made a lower average error score on the sum total of the discrimination-reversal problems than that made by the normal monkey (subject 52) with the poorest performance.

These data suggest that differences in performance between normal and prefrontal monkeys exist in discrimination learning and discrimination-reversal. Since this does not result from an inability on the part of the operated animals to solve such problems, it might be explained either as a general decrease in ability to adapt to new situations, or it might result from loss within some more specific function which acts to facilitate discrimination learning and discrimination-reversal.

VII. Summary and Conclusions

(1) Six normal monkeys and 4 monkeys with bilateral prefrontal lobectomies learned a series of 10 discrimination problems and 10 discrimination reversal problems.

(2) The normal monkeys learned both problems with fewer errors than the operated animals and the differences approached statistical significance.

(Manuscript received October 22, 1942)

References

(1) Fisher, R. A. *Statistical methods for research workers.* London: Oliver and Boyd, 1938.
(2) Harlow, H. F. Responses by rhesus monkeys to stimuli having multiple sign-values. In McNemar, Q., & Merrill, M. A., edit., *Studies in Personality.* McGraw-Hill Book Co., 1942, 105–123.
(3) Harlow, H. F. Studies in discrimination learning by monkeys: I. The learning of discrimination series and the reversal of discrimination series. *J. gen. psychol.* (In press).
(4) Harlow, H. F., & Bromer, J. A. A test apparatus for monkeys. *Psychol. Rec.*, 1938, 2, 434–436.
(5) Klüver, H. Certain effects of lesions of the occipital lobes in macaques. *J. Psychol.*, 1937, 4, 383–401.

[3] Furthermore, prefrontal subjects, number 54 and 55, averaged only 1.5 errors on an additional set of 20 discrimination-reversals, clearly showing the ability of prefrontal monkeys to master problems of this type if given additional training.

BILATERAL ELECTROENCEPHALOGRAMS FROM NORMAL SPEAKERS AND STUTTERERS[1]

BY JOHN R. KNOTT AND THEODORE D. TJOSSEM

University of Iowa

Introduction

On the basis of certain clinical and experimental evidence, Travis (6) has advanced a theory of stuttering based upon the allied concepts of vertical and lateral dominance in the cerebro-spinal axis. According to this theory, stuttering is a neuromuscular dysfunction associated with lack of a sufficiently dominant lateral gradient, such a gradient being considered essential to normal integration of speech functions.

The dominant lateral gradient is said to be a center of highest metabolic activity and excitation which (in a right-handed individual) lies in the left cerebral hemisphere. This dominant center "furnishes an inhibitory directional control" (6, p. 34) over contralateral as well as over lower centers.

Douglass (1), who recorded bilateral EEGs from 20 stutterers and 20 normal speakers under conditions of silence and speech, attempted to find differences between the two groups with respect to inter-hemisphere differences in percent time alpha present. His results showed a meaningful division between the groups when measuring the difference between the two hemispheres in percent time alpha present during silence and thus gave the first fairly direct measure of this lateral gradient and its variation in stutterers. Stutterers, as a group, tended to have a higher percent time alpha present in the left than in the right occipital area, while non-stutterers tended to have a lower percent time alpha present in the left than in the right occipital area. These data are of especial interest in the light of Raney's findings (7) that in identical twins the percent time alpha present was greater in the non-dominant than in the dominant hemisphere (when the dominant hemisphere was defined on the basis of the phi-test).

These data of Douglass appear to be of even greater significance when it is realized that the differentiation occurred during silence and not during speech, for the dominance theory postulates a 'pathophysiological subsoil' (6, p. 14) upon which stuttering is built.

[1] From the Laboratory of Physiological Psychology, Department of Psychology, and the Electroencephalographic Laboratory, Department of Psychiatry, University of Iowa.

Statement of Problem

The purposes of this investigation were: to extend the observations of Douglass by the addition of further samples selected from the population of stutterers; to discover whether or not these samples, together with the Douglass sample, formed a homogeneous distribution; and to discover whether this distribution differed significantly from the Douglass sample of non-stutterers.

Apparatus and Procedure

Electroencephalograms were recorded simultaneously from the left occipital (LO), right occipital (RO), left motor (LM), and right motor (RM) areas. (In one group of subjects records were taken simultaneously from the LO and RO areas, and a few minutes later were taken simultaneously from the LM and RM areas.)

All leads were monopolar, with reference ('ground') leads attached in parallel to the lobes of the ears. The motor area leads were placed at a point 3 cm. from the mid-line along a line approximating the central fissure, determined by Chiene's method. The occipital area leads were placed 3 cm. on either side of the mid-line and at a level 3 cm. above the inion.

The equipment consisted of 4-stage resistance-capacitance-coupled amplifiers with a time constant of 0.25 sec., and a Grass 4-element inkwriting oscillograph.

The subjects were seated in a comfortable chair in a semi-soundproofed, darkened, electrically shielded room. They were instructed to sit quietly with eyes closed while the record was taken.

Subjects

The subjects used in this investigation came from two sources and are consequently treated as two separate groups. The first group was composed of 24 stutterers enrolled in the Speech Clinic of the State University of Iowa. Twenty-one were males; 3 were females. The age range was from 11 to 32 years; 3 were below 16 years of age. The mean age was 20.8 years. The second group was composed of 14 adult stutterers enrolled in the Speech Clinic of the University of Minnesota. Eleven were males; 3 were females.

Supplemental data have been taken from the investigation of Douglass. These are based upon a sample of 20 adult stutterers enrolled in the Speech Clinic of the State University of Iowa and 20 adult normal speakers.

Analysis of Data

Analysis of the data consisted of reading in sequence 5 meters of record to determine the amount of alpha rhythm present per meter of record from each cortical area. Alpha rhythm was defined as 3 or more consecutive waves, falling within a frequency range of 8.5 to 12 per sec. In the event of an artifact being present in one or more of the leads, that portion of the record was deleted and an equivalent length of record added at the end of the meter in which the artifact occurred.

The percent time alpha present (cm. alpha per meter/100 cm.) was computed for each area. The difference in percent time alpha present in contralateral areas was computed by subtracting percent time alpha present on the right from percent time alpha present on the left.

The reliability of reading was checked by re-reading 3 meters of unmarked record after an interval of one week. There was 94 percent agreement between the two readings.

Results

The statistical analysis was concerned with a comparison between the two groups of stutterers of this study and the group of stutterers and the group of non-stutterers reported by Douglass. Comparisons

were made among these four groups with respect to LO minus RO differences and to LM minus RM differences.

The differences in the interhemisphere distribution of percent time alpha in the occipital and the motor areas are presented in Tables 1 and 2.

TABLE 1

MEAN LO MINUS RO PERCENT TIME ALPHA DIFFERENCES FOR STUTTERERS AND NORMAL SPEAKERS

Group	N	AM	Range
Iowa Stutterers	24	+0.57	+17.52 to −11.26
Minn. Stutterers	14	−0.98	+22.09 to −14.89
Douglass Stutterers	20	+1.91	+12.08 to −10.47
Total for Stutterers	58	+0.66	+22.09 to −14.89
Douglass Normals	20	−3.61	+11.55 to −20.80

TABLE 2

MEAN LM MINUS RM PERCENT TIME ALPHA DIFFERENCES FOR STUTTERERS AND NORMAL SPEAKERS

Group	N	AM	Range
Iowa Stutterers	24	+2.00	+11.64 to −10.88
Minn. Stutterers	14	+0.36	+ 4.37 to − 6.34
Douglass Stutterers	13	−1.45	+ 4.96 to −15.90
Total for Stutterers	58	+0.4219	+11.64 to −15.90
Douglass Normals	20	+0.55	+11.18 to −12.14

The data pertaining to the occipital area were first analyzed in terms of the three groups of stutterers, using an analysis of variance. The between groups variance was 34.54; the within group variance was 48.34; F was thus equal to 0.71. This F is not significant, indicating that the three groups of stutterers were not drawn from different populations. That is, the three samples form a homogeneous group.

Treating the three groups of stutterers as a single sample of randomly selected groups, it was possible to compare the mean LO minus RO difference with the mean LO minus RO difference as obtained in the group of normal speakers reported by Douglass. A t of 2.52 was obtained, which indicates the difference (4.27) to be significant at the two percent level of confidence.

Inspection of individual LO minus RO differences indicated that 5 of Douglass' 20 normals (25 percent) had positive scores; 12 out of his 20 stutterers (60 percent) had positive scores. In the two samples collected in this investigation, one, the Minnesota group, 5 of 14 differences were positive (36 percent); in the other, 10 out of 24

differences were in the positive direction (42 percent). Thus, there is a tendency for groups of stutterers to have a greater proportion of positive differences in LO minus RO scores than occurs in a group of non-stutterers. There is, however, a great overlapping of the distribution of actual differences, which may be seen in Table 1.

Similar treatment was given the LM minus RM data. Mean values of these differences are presented in Table 2. Comparison of the three groups of stutterers by an analysis of variance yielded an F equal to 3.20. (Between groups variance equalled 64.48; within group variance equalled 20.18.) This value of F indicates significance at the five percent level of confidence. Apparently, with relation to these scores, the groups do not represent a homogeneous distribution. When the difference between the mean scores for the Iowa, Minnesota, and Douglass groups were compared, however, only one difference proved to reach significance, that obtaining between the Iowa and the Douglass stutterers. In this instance t was 2.53, indicating significance at the two percent level of confidence. However, an F indicating significance at the five percent level is at a sufficiently low level of confidence to suggest that such a level has been reached by chance and that the true differences are zero. Since the supposed significance is reached in only one of three possible comparisons, this hypothesis seems plausible.

When LM minus RM differences were evaluated for the three groups of stutterers considered as a single sample and when this value compared with that for the Douglass normals, the difference was 0.1 and was obviously not significant.

Discussion

The data presented above indicate that the two new groups of stutterers do not differ significantly from the group of stutterers studied by Douglass, and this finding tends to corroborate his data. It also serves to strengthen his conclusion that, during silence, stutterers and normals, considered as groups, can be electroencephalographically differentiated in terms of the distribution of differences in left and right occipital area percent time alpha. In terms of the distribution of left and right motor area percent time alpha, which Douglass found to be non-differentiating with respect to stutterers and normal speakers, the present data again lend support. Consequently, one can feel somewhat more certain that the results of the Douglass study were not determined by chance variation in the samples and that a real difference exists between some of the electrocortical characteristics of stutterers and normal speakers.

Jasper (3) has suggested that the percent time alpha may be related to what he terms 'cortical excitatory state.' This suggestion

is based upon the empirical fact that stimulation, which should lead to a greater excitatory state, decreases the percent time alpha present. It is obvious, of course, that the term 'central excitatory state' is still not empirically demonstrated but is inferred.

The hypothesis suggested by Jasper, that a decrease in alpha represents an increase in cortical excitatory state, could be applied to the problem here investigated and would serve to 'explain' the differences between stutterers and non-stutterers in terms of lateral differences in cortical excitatory state. That hemisphere showing the least alpha would thus be said to be at a higher level of excitation than the other hemisphere. Thus, it may be that lateral differences in percent time alpha measure central laterality. If so, the difference in the central laterality of stutterers and normal speakers postulated by the dominance theory has been experimentally demonstrated.

Further verification of these findings, as well as of hypotheses used in their explanation, is essential. Of particular importance is the continued study of groups of non-stutterers for, since we now know that stutterers form a homogeneous group, it is necessary to know whether or not repeated samples of normal speakers will reveal a homogeneous group different from stutterers. On the basis of this report it would be *expected* that they would; this prediction must be tested, however.

In the meantime, one can conclude that there are possible avenues of experimental differentiation between certain electrocortical characteristics of stutterers and non-stutterers; that these are fundamentally physiological in their expression and are thus assumed to be physiological in their origin; and that these differences must be accounted for by any adequate theory of stuttering.

Summary and Conclusions

(1) Douglass has recently reported data which might be interpreted as offering support to the lateral dominance theory of stuttering.

(2) This study endeavored to confirm the Douglass study. The procedure followed involved the addition of two new groups of stutterers to Douglass' data in order to further test his conclusion that with regard to the bilateral distribution of the alpha rhythm, stutterers form a homogeneous distribution different from that formed by normal speakers.

(3) The results indicate that when these comparisons are made, the stutterers form a homogeneous distribution with respect to differences in alpha index in the left and right occipital areas (L minus R), and that when compared with the non-stutterers, the L minus R scores are significantly lower (more positive) than those

scores for non-stutterers. No such differences appeared in the data from the motor area.

(4) These findings may be tentatively evaluated in terms of Jasper's hypothesis of cortical excitatory state. Such an evaluation would lead to the conclusion that stutterers tend to have a different lateral (hemispheric) excitability than do non-stutterers; this is powerful support for that theory of stuttering based on the concept of lateral cerebral dominance.

(5) There is a difference, however, between the statistical average of stutterers (or non-stutterers) and any given stutterer (or non-stutterer). The overlap in the distribution of L minus R occipital percent time alpha will not permit generalization to the effect that a difference in cortical excitatory state is the only etiological factor related to stuttering.

(6) Further study of this problem is to be encouraged.

(Manuscript received September 19, 1942)

References

1. Douglass, L. C. A study of bilaterally recorded electroencephalograms of adult stutterers. Unpublished Doctor's Dissertation, Univ. Iowa, 1940.
2. Hamstra, R. H. Central and peripheral rhythms: their relation to moments of stuttering. Unpublished Master's Thesis, Univ. Iowa, 1939.
3. Jasper, H. H. Electrical signs of cortical activity. *Psychol. Bull.*, 1937, **34**, 411–481.
4. Lindsley, D. B. Bilateral differences in brain potentials from the two cerebral hemispheres in relation to laterality and stuttering. *J. exp. Psychol.*, 1940, **26**, 211–225.
5. Scarbrough, H. E. A quantitative and qualitative analysis of the electroencephalogram of stutterers and non-stutterers. Unpublished Doctor's Dissertation, Univ. Iowa, 1939.
6. Travis, L. E. *Speech Pathology*. New York: Appleton, 1931.
7. Raney, E. T. Brain potentials and lateral dominance in identical twins. *J. exp. Psychol.*, 1939, **21**, 24–39.
8. Travis, L. E., & Knott, J. R. Brain potentials from normal speakers and stutterers. *J. Psychol.*, 1936, **2**, 137–150.
9. Travis, L. E., & Knott, J. R. Bilaterally recorded brain potentials from normal speakers and stutterers. *J. Speech Disorders*, 1935, **2**, 239–241.